Analyzing gene function in the mouse: gene trap mutations in *Arp3*, *Lsm16*, and a novel KRAB-ZF gene

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr.rer.nat.)

genehmigte
D i s s e r t a t i o n

von Blair Raymond Prochnow
aus St. Cloud, MN, USA

1. Referent: Professor Dr. Hans-Henning Arnold
2. Referentin: Professorin Dr. Brigitte M. Jockusch
eingereicht am: 15.05.2006
mündliche Prüfung (Disputation) am: 28.07.2006

2007

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor oder Betreuer der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Vauti, F., Ramasamy, S. K., Goller, T., Beine, R., Prochnow, B., Meyer, N., Ruiz, P. & Arnold, H.-H.: Establishment of mutant animal models for human diseases on the basis of a gene trap approach in mouse embryonic stem cells. (Poster) Human Genome Meeting, Braunschweig (2001).

Vauti, F., Prochnow, B. & Arnold, H.-H.: Analysis of a new mouse line with a gene trap vector integration in a KRAB-zinc finger encoding gene. (Poster) 17th International Mouse Genome Conference, Braunschweig (2003).

Table of Contents

Summary

Elucidation of the function of the estimated 30,000 genes in the mouse genome is an important issue in molecular genetics. Of special interest are those genes of biomedical relevance to human diseases. One of the most informative means for studying gene function in the mouse is by analyzing mutants. A well-established large-scale approach, which allows for the mutation and analysis of novel, as well as known genes, is the gene trap approach. The work in this thesis involves the analysis of three gene trap mouse lines.

The first line represents a gene trap mutation in the *Arp3* gene. Arp3 is a component of the Arp2/3 complex, which is involved in actin polymerization and branching. The *Arp3* mutant mouse has shown that the mutation is lethal in homozygous mutant embryos around stage E3.5, demonstrating an essential role of Arp3 in early embryogenesis. Mutant embryos up to the early blastocyst stage appear normal, most likely because they contain maternal Arp3 protein. However, mutant blastocysts fail to fully expand and are arrested for further development. In heterozygous mice, the levels of Arp3 transcripts and protein are reduced by approximately 50%, but no apparent phenotype was detected. Hence, the mutation is recessive lethal, with no gene dosage effect in heterozygous mice.

In the second mutant mouse line, a novel krüppel-associated box zinc-finger gene (KRAB-ZF) has been trapped. Members of the KRAB-ZF family act as transcriptional repressors. The trapped gene is expressed in two splice variants throughout the mouse at all stages of life, though expression is not found in every cell. Beyond hypopigmentation, which was found to result from mutations carried by the original ES-cells, no obvious phenotype could be detected in mutant mice. This indicates that the role of the gene is of only limited importance, probably due to redundant genes.

The third line analyzed contains gene trap vector integration in a novel gene. Mice homozygous for the trapped gene display a general reduction in body size and weight, which persists from juvenility through adulthood. This reduction is found in all tissues analyzed, but most notably in adipose tissues. In muscle and adipose tissues, the fusion protein of the trapped gene appears expressed in cytoplasmic foci. In the

case of muscle, the fusion protein is differentially expressed, with all muscle cells of the soleus muscle showing expression, whereas only few cells of the gastrocnemius muscle exhibit expression. Furthermore, in tissues from adult mutants, altered expression of several adipocyte and energy metabolism marker genes was found. While finalizing this thesis, the novel gene could be identified as the mouse ortholog to the human *Lsm16* and yeast *Edc3* genes. In mouse, this gene has therefore been named *Lsm16*. In human and yeast, these genes are only recently known to be involved in enhancing mRNA decapping. Since the mouse Lsm16 protein shares over 95% identity to its human homolog, it is implied here that the mouse protein is also involved in mRNA decapping.

1. Introduction

With the successful completion of the Human Genome Project, a powerful resource for enhancing biological research was introduced. What once was the goal of generating and aligning raw genomic sequences has now become the utilization of this information to search for and identify genes. Here, of great interest are genes of biomedical importance.

To characterize and analyze the function of known and novel genes, an appropriate model system is vital. Due to ethical, legal and logistical restrictions, research on humans is limited; therefore, more appropriate and acceptable systems have been established. Among the most commonly used eukaryotes are *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Xenopus laevis*, *Danio rerio* and *Mus musculus*. Although each of these organisms has proven its importance, each is limited by its own set of advantages and disadvantages. The choice of which model organism is most appropriate is influenced by such factors as the actual or expected role of the gene, method of analysis, as well as the process, organ or cell type to be analyzed. In the case of each model system, the analysis of mutants is of monumental importance in the ability to elucidate gene function.

The mouse, with its genome having been published in 2002, represents the common model organism most closely related to humans, especially if the intention is to investigate the role of genes in the whole mammal. This makes the mouse highly relevant for the establishment of disease models of human genetic disorders. Advantages of the mouse over other mammalian systems include small size, ease of maintenance in the laboratory, a short breeding cycle and ease of mutation. Here, mutations can be introduced by irradiation, treatment with chemical mutagens or insertion of DNA fragments into the genome to interrupt genes.

In the case of irradiation or chemical mutagens, the genome is mutated randomly and the identification of the mutated locus can be time consuming. Mutagenesis through the insertion of DNA into the genome though, such as in gene targeting, involves the mutation of a known locus using a constructed sequence. An approach that includes desirable attributes of both these methods is gene trapping. This

technique consists of creating random mutations, thus known as well as novel or uncharacterized genes can be mutated, but because the mutation is created by insertion of a constructed sequence, the locus carrying the mutation can be easily determined.

1.1. The gene trap approach in the mouse

The gene trap approach, with respect to the mutant mouse lines analyzed in this thesis, is based on the creation of a mutation at a random locus in the genome of mouse embryonic stem (ES) cells by integration of the splice-acceptor vector PT1βgeo or PT1ATGβgeo (Fig. 1). These vectors are 8.9 kb in size and contain four important sequences in their *βgeo* functional cassette (Gossler *et al*., 1989; Friedrich and Soriano, 1991; Skarnes *et al*., 1992). The first sequence of importance is the splice-acceptor site (SA) as part of the *engrailed2* intron and exon. This sequence ensures that splicing of the endogenous transcript includes the functional cassette of the vector. The second sequence, which is fused in-frame, is the *β-galactosidase* (*lacZ*) reporter gene. Expression of the *lacZ* gene enables demonstration of where and when the trapped gene is expressed. The third sequence is the *neomycin phosphotransferase* (*neo*) gene. This is a selection marker of use during the incubation of the ES-cells. Only clones with integration of the promoterless splice-acceptor vector in an actively transcribed gene will survive in selective medium containing the neomycin analog G418. The fourth sequence in the cassette is the simian virus polyadenylation signal (SVpA) at the 3´ end.

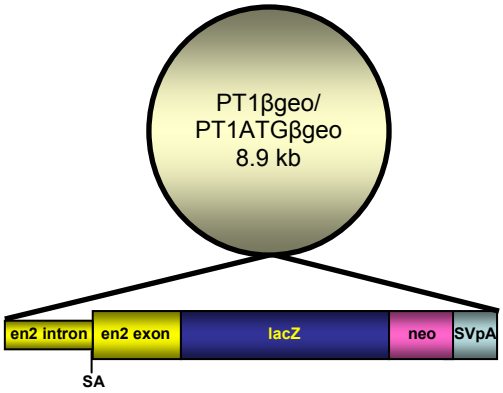


Figure 1. Schematic diagram of the 8.9 kb PT1βgeo/PT1ATGβgeo splice-acceptor gene trap vector. Expanded in detail is the *βgeo* functional cassette. en2: *engrailed2*, SA: splice-acceptor site, lacZ: *β-galactosidase* gene, neo: *neomycin phosphotransferase* gene, SVpA: simian virus polyadenylation signal.

The PT1ATGβgeo vector differs from the PT1βgeo vector in that it contains an ATG translation initiation codon near the 5´ end of the *βgeo* cassette. This allows translation of the mRNA even if the vector has integrated 5´ of the endogenous translation initiation codon. The final vector constructs of PT1βgeo and PT1ATGβgeo have been cloned by Dr. Franz Vauti for use in the German Gene Trap Consortium.

During splicing of the transcribed gene, the splice-acceptor in the *engrailed2* sequence ensures that the *lacZ/neo* cassette (*βgeo*) becomes fused in-frame to the endogenous exon lying immediately 5´ to the site of vector integration. Using the 5´ RACE (Rapid Amplification of cDNA Ends) method, the mRNA sequence 5´ to the *βgeo* cassette can be determined. This is referred to as the trapped sequence. This sequence can be utilized to acquire more information pertaining to the trapped locus. Following transcription and splicing, the mRNA from the trapped gene is lacking the endogenous sequence 3´ to the vector. Upon translation, a fusion protein is formed which contains the amino acid sequence encoded by the endogenous exons 5´ to the vector integration site, as well as part of the *engrailed2* exon, and the *lacZ* and *neo* amino acid sequences.

Because the gene trap approach is based on random vector integration, known, as well as unknown genes can be trapped. For this thesis, three gene trap mutant mouse lines were selected for analysis. Each of these lines represents a variation on the level to which the trapped locus has already been characterized. The mutant line A009F03 carries a trapped gene that is known and well characterized. The A20010 line contains a mutation in an unknown gene, but whose sequence has allowed for the precise classification of the gene. In the A008A01 line, the trapped gene is unknown with no well-classified domains.

1.2. Gene trap vector integration in the Arp3-encoding gene

The first mutant mouse line analyzed in this thesis is designated A009F03 and displays insertional mutation in a known, well-characterized gene. Integration of the PT1βgeo vector is in the locus encoding *actin-related protein 3* (*Arp3*).

1.2.1. Arp3 protein is a subunit of the Arp2/3 complex

Arp3, which derives its name from its homology (~45% identity) to actin, is one of the subunits in the seven-subunit Arp2/3 protein complex (Robinson *et al*., 2001). This complex has been found in a diversity of organisms ranging from *Acanthamoeba castellanii* where it was first discovered (Machesky *et al*., 1994) to humans (Welch *et al*., 1997) and higher plants (Le *et al*., 2003).

The Arp2/3 complex is known to localize to areas of dynamic actin cytoskeleton activity, where it plays a pivotal role in regulating actin filament formation by initiating *de novo* nucleation of filaments from monomeric actin and cross-linking them into a branching network (Mullins *et al*., 1998; Pollard and Beltzner, 2002). This is found at the leading edge of lamellipodia in animal cells and in motile actin patches in yeasts (Machesky *et al*., 1997; Welch *et al*., 1997; Moreau *et al*., 1996). There are also several pathogens that utilize the host's actin cytoskeleton to move intra- and intercellularly. These include the bacteria *Listeria monocytogenes*, *Shigella flexneri* and *Rickettsia* species, as well as enteropathogenic *Escherichia coli* and the Vaccinia virus (reviewed in Gouin *et al*., 2005). Additionally, endosomes, lysosomes, pinocytic vesicles and mitochondria have been shown to utilize the Arp2/3 complex and the actin cytoskeleton for intracellular motility (Taunton *et al*., 2000; Engqvist-Goldstein and Drubin, 2003; Boldogh *et al*., 2001). In plants, the Arp2/3 complex is involved in actin bundle organization leading to proper epidermal cell shape (Le *et al*., 2003).

The activity of the Arp2/3 complex and its capability to nucleate actin filament formation are regulated by nucleation-promoting factors (NPFs) such as members of the WASp/Scar family, which are in turn regulated by several signaling pathways (Machesky *et al*., 1999; Welch and Mullins, 2002).

The exact role of the Arp3 subunit is still unclear, but it has been shown that Arp3 binds ATP, which leads to a rearrangement of the Arp2/3 complex. It is proposed that this conformation favors nucleation of new actin filaments (Robinson *et al*., 2001).

1.2.2. Aims of this work for the A009F03 mutant mouse line

In a preliminary analysis of the *Arp3* mutant line, it was found that the *Arp3* mutation was recessive embryonic lethal (Diploma thesis: Freese, 2003). Therefore, the aims

of this work were to define the time and likely cause of embryonic lethality. For this, embryos from heterozygous matings were isolated at day 3.5 postcoitum (pc) and cultured overnight to study their morphology and development up to lethality. In addition, day 3.5 pc embryos were analyzed for expression of the Arp3 protein.

1.3. Gene trap vector integration in a KRAB-ZFP coding gene

The mutant mouse line designated A20010 shows vector integration in a novel gene, which encodes a krüppel-associated box zinc-finger protein (KRAB-ZFP).

1.3.1. Krüppel-associated box zinc-finger proteins

KRAB-ZFPs belong to the family of krüppel-type proteins defined by a variable number of C_2H_2 zinc-finger motifs at the C-terminus. These motifs consist of a short amino acid sequence with two cysteine and two histidine residues that bind to a zinc ion and fold into a finger-like structure. Consecutive repeats of this motif establish the basis for interaction with DNA. There are an estimated 600 to 700 genes that contain C_2H_2 zinc-finger motifs and act as transcription factors in the human and mouse genomes (Klug and Schwabe, 1995).

Approximately one-third of the krüppel-type zinc-finger genes also possess a krüppel-associated box (KRAB) as effector domain (Bellefroid *et al*., 1991). The KRAB domain consists of approximately 50-75 amino acid residues and is divided into A and B boxes. The A box has been noted to be a DNA binding-dependent transcription repression module, which carries out this function by binding corepressors, whereas the B box enhances the repression mediated by the A box (Margolin *et al*., 1994; Witzgall *et al*., 1994). This repression is at least partially mediated by the binding of Transcription Intermediary Factor 1β (TIF1β), also known as Transcriptional Corepressor KRAB-Associated Protein-1 (KAP-1), or KRAB-A Interacting Protein (KRIP-1) (Agata *et al*., 1999; Friedman *et al*., 1996; Kim *et al*., 1996; Moosmann *et al*., 1996). The current model of transcriptional repression by KRAB-ZFPs suggests that the protein binds to the promoter region of the target gene via the zinc-finger motifs, followed by TIF1β binding as an oligomer to the KRAB domain. TIF1β then recruits Heterochromatin Protein 1 (HP1), histone deacetylases (HDACs) and SET-Domain Binding Protein 1 (SETDB1) to the complex (Urrutia, 2003). SETDB1 is a methyltransferase that methylates lysine-9 of histone H3

(Schultz *et al*., 2002). This silences gene expression by leading to heterochromatin formation at the target promoter.

KRAB zinc-finger proteins have been described primarily in higher vertebrate species. From the KRAB-ZFPs for which specific functions have been postulated, no information has emerged that would allow grouping these proteins into specific functional categories. KRAB-ZFPs have been shown to participate in a wide spectrum of biological processes in humans and mice including cardiac development (Dai and Liew, 1999), osteoblastic differentiation (Jheon *et al*., 2001), neuronal functional maturation (Li *et al*., 2002), liver fibrogenesis (Schnabl *et al*., 2005), regulation of lineage commitment during hematopoietic development (Mark *et al*., 2001), as well as cell proliferation, apoptosis and repression of RNA polymerase II (reviewed in Urrutia, 2003). The only KRAB-ZFPs where the biological function has been elucidated are the *Regulators of sex-limitation* alleles *Rsl1* and *Rsl2*, which accentuate sex differences in hepatic gene expression in the mouse (Krebs *et al*., 2005).

1.3.2. Aims of this work for the A20010 mutant mouse line

During the preliminary characterization of the mouse line, hypopigmentation was observed in homozygous mutants (Diploma thesis: Prochnow, 2002). The aims of this thesis work were, therefore, to elucidate the cause of hypopigmentation in mutant mice. This involved the statistical verification of the correlation between coat color and genotype, characterization of the pigmented tissues, i.e. eyes and coat, analysis of the expression of the KRAB/β-galactosidase fusion protein in embryos and adult organs, and the examination of genes involved in pigmentation or development of pigment producing cells in embryos and organs.

1.4. Gene trap vector integration in a novel gene containing a YjeF-related protein N-domain

The mutant mouse line A008A01 carries a novel trapped gene that contains no classified domains except for a proposed domain at its C-terminus with similarity to the N-terminus of the YjeF protein found in *Escherichia coli*. This domain in the A008A01 line is referred to as a YjeF-related protein N-terminus domain (Yjef-N).

1.4.1. The YjeF protein

Although the YjeF protein in *E. coli* is largely uncharacterized, the Protein families database of alignments (Pfam) categorizes the protein according to its structure as a member of the carbohydrate kinase family. This family belongs to the ribokinase-like superfamily. The domain proposed as responsible for ribokinase activity is found at the C-terminus. The function of the N-terminal domain, YjeF-N, has not yet been determined (Anantharaman and Aravind, 2004).

1.4.2. Aims of this work for the A008A01 mutant mouse line

In the initial analysis of the A008A01 line by Nicole Meyer (Diploma thesis: Meyer, 2003) and Fabian Garreis (Diploma thesis: Garreis, 2005), it was observed that most of the homozygous mutant mice were reduced in weight and some demonstrated a reduced “latency to fall” in the Hanging-Wire Test, which might be caused by muscle weakness. Additional information to structure and function of the YjeF-N containing protein was not available at the time. The aims of this thesis work with respect to the A008A01 mouse line were focused on three aspects. First, a more detailed analysis of the reduced body weight in mutants, second, a histological evaluation of leg muscle, and third, continued bioinformatic analysis of the trapped gene. A reduction of body weight in mutants has already been established, as mentioned above. To determine the precise cause for this reduction, the various tissues and organs of wild-type and homozygous mutant mice will be compared. Special attention will be paid to the analysis of fat tissue, as well as skeletal muscle of the legs. Additionally, because of the continuous update of bioinformation databases, as well as the constant improvement to the knowledge base of characterized genes and domains, it is a goal of this thesis to use these resources to better characterize the protein sequence of the A008A01 trapped gene and possibly assign a function to it.

2. Materials and Methods

2.1. Materials

2.1.1. Chemicals, reagents and enzymes

Chemical and reagents were of p.A. quality and purchased from Fluka, J.T. Baker, Merck, Roth and Sigma. The radioisotope α^{32}P-dCTP was from Amersham. Enzymes were bought from New England Biolabs, Qiagen, Amersham and Invitrogen.

2.1.2. Labware

Labware was purchased from Becton Dickinson, Eppendorf, Greiner, Nunc, Sarstedt and Menzel-Gläser, unless stated otherwise.

2.1.3. Oligonucleotides

The oligonucleotides utilized as primers for genotyping, sequencing and KRAB-ZF specific RT-PCR (Tab. 1), as well as for detecting the expression of pigmentation genes in the KRAB-ZF mutant mouse line (Tab. 2) were designed using Primer Select (DNASTAR). The primers for adipocyte and energy metabolism genes in the *Lsm16* gene trap line (Tab. 3) were taken from Harada *et al*., (2003) unless specified otherwise. The *Rpl7* primers used as quantity controls are included in table 2. All oligonucleotides were synthesized by Roth.

Table 1. Primers for genotyping, sequencing and KRAB-ZF specific RT-PCR.

Name	Purpose	Sequence 5´ to 3´
Arp3 U1	genotyping *Arp3* locus	AACCCACAACAACAAAAAA
Arp3 L1	genotyping *Arp3* locus	TCCCCTCCTCTTACAACAC
LacZ L2	genotyping *Arp3* locus	TGGGCAAGAACATAAAGTGA
LacZ U1	genotyping KRAB-ZF locus	CGCCATTTGACCACTACC
LacZ L1	genotyping KRAB-ZF locus	GGTGGCGCTGGATGGTAA
Lsm16 U1	genotyping *Lsm16* locus	AGCCCAGCAGCCTCTTAC
Lsm16 L1	genotyping *Lsm16* locus	GCCTGGCTCCCTCTACAT
En2 in L1	genotyping *Lsm16* locus	CAACCTCCGCAAACTCCT
p U2*	genotyping *p* locus	CACTATGCCTGCTTGGATGCTGC
p L2*	genotyping *p* locus	ATGGGACATCTGTGGTTACACTGG
KRAB-ZF U1	RT-PCR for KRAB-ZF transcripts	TACCGATGTCTGAGTTGTCC
KRAB-ZF L1	RT-PCR for KRAB-ZF transcripts	CTGTGGTAATGGACTTCA
KRAB-ZF U2	Sequencing KRAB-ZF PCR fragments	TACCGATGTCTGAGTTGTCC
KRAB-ZF U3	Sequencing KRAB-ZF PCR fragments	AGCGGACGTTGTATAGAG
KRAB-ZF L2	Sequencing KRAB-ZF PCR fragments	ATTCTGCTATGAGTTTCA
KRAB-ZF L3	Sequencing KRAB-ZF PCR fragments	GGTTTTAAGAGTTGAGCC
KRAB-ZF L4	Sequencing KRAB-ZF PCR fragments	CTGTGGTAATGGACTTCA
Tyr U2	Sequencing to determine *Tyr* allele	ACCGCCCTCTTTTGGAAGTT
Tyr L2	Sequencing to determine *Tyr* allele	CTCAATTAGTTGTAAGAGGA

* sequence taken from Brilliant *et al*., 1994

Table 2. Primers for pigmentation genes analyzed in the KRAB-ZF mutant mouse line.

Name	Sequence 5´ to 3´	Name	Sequence 5´ to 3´
A U	TGCGGAATAGAGTCACTTGT	A L	CCCACAAGTCACAACCACTG
Atrn U	CTGGCTCTTCTGGATTTGTA	Atrn L	CAGTGAGGAAAGCAGGAGCA
Dct U	TGGGGGCTTCTGCTGGGTTG	Dct L	CCACACAAAAAAGTCATACA
Ednrb U	ATGGGGAGAGAAAAGAGGAT	Ednrb L	CAGAACCACAGAGACCACCC
Ggt1 U	TGGCTGTGGTTCTGGTATTT	Ggt1 L	GTTTTTTGTCCAGGGCTATT
Kit U	CGCCATCCATCCATCCAGCA	Kit L	TGTAGACACATCTTTTATGG
Mc1r U	GGGCTGGTGAGTCTGGTGGA	Mc1r L	CCCCAGCACAGGAAGAAAAT
Mgrn U	CCCCCATGAGCCAGTGAAGA	Mgrn L	GTCCACAATTTGCTTCTGCT
Mitf U	ACCCCACCAAGTACCACATA	Mitf L	CTTGAGAATGGTTCCCTTGT
Ostm1 U	TCGCAATCAAGATGGACAAC	Ostm1 L	GCAGGTGACCGAACAGTTGA
p U	CTGCCATCCTTTCTGCCTTC	p L	CAGAATAGCAATCCATCCAA
Pax3 U	CGCCCCGCCTCCCTCTCTGG	Pax3 L	GGGTTCTCTCTTTTGTATTC
Pomc1 U	AGGCGACGGAAGAGAAAAGA	Pomc1 L	GTGGACTCGGCTCTGGACTG
Si U	TCCCGTGCTTGTGCTGAGTG	Si L	CCAGGTCTTCCAAACATAAA
Slc45a2 U	TGGCGTGGAAGGTATTCTGA	Slc45a2 L	CAGAAGGATAGGTCAGGGTA
Tyr U1	CACCCTGAAAATCCTAACTT	Tyr L1	CAGCCATTGTTCAAAAATAC
Tyrp1 U	CCCTTTTCCTGATGCTGTTT	Tyrp1 L	GAGGAAGGTTTTTTTGACTG
Rpl7 U	GAAGCTCATCTATGAGAAGGC	Rpl7 L	AAGACGAAAGGAGCTGCA

Table 3. Primers for adipocyte and energy metabolism genes analyzed in the *Lsm16* mutant mouse line.

Name	Sequence 5´ to 3´	Name	Sequence 5´ to 3´
Acly U	TGGAGGCAGCATTGCAAAC	Acly L	CGCTGGCATTAAGGAGGAAG
Adipoq U	CGGCAGCACTGGCAAGTT	Adipoq L*	GTCCCCATCCCCATACACCT
Adn U*	CCTGAACCCTACAAGCGATG	Adn L	GCGCAGATTGCAGGTTGTC
Dgat1 U	CATGCGTGATTATTGCATCCA	Dgat1 L	GCCAGGCGCTTCTCAATCT
Dgat2 U*	ATTTGGCTACGTTGGCTGGT	Dgat2 L	AGGAATAAGTGGGAACCAGATCAG
Fasn U	TCCTGGAACGAGAACACGATCT	Fasn L	GAGACGTGTCACTCCTGGACTTG
Gpam U	CAACACCATCCCCGACATC	Gpam L*	AGGGGGGCAGATACAGGTTT
Hmgcr U	GACAAGAAGCCTGCTGCCATA	Hmgcr L	CGGCTTCACAAACCACAGTCT
Hmgcs1 U	TGCACGGATCGTGAAGACA	Hmgcs1 L	GTCTCTCCATCAGTTTCTGAACCA
Hsl U	CTCCTCATGGCTCAACTCCTTCC	Hsl L	AGGGGTTCTTGACTATGGGTG
Insr U	CGCTGTGTGAACTTCAGCTTCT	Insr L	CAGCCAGGCTTCCGAGAGT
Irs1 U	CCTCAGTCCCAACCATAACCA	Irs2 L	CCGGCACCCTTGAGTGTCT
Ldlr U*	TGAGTGCTGTGTCCCAGC	Ldlr L	GCGGTCCAGGGTCATCTTC
Lep U	AACCCTCATCAAGACCATTGTCA	Lep L	CCTCTGCTTGGCGGATACC
Lpl U	AGTGGCCGAGAGCGAGAAC	Lpl L	CCACCTCCGTGTAAATCAAGAAG
Pparg U	GCCCACCAACTTCGGAATC	Pparg L	TGCGAGTGGTCTTCCATCAC
Retn U	TGGACCTTGGCAGGACTG A	Retn L	TCTCCTTCCACCATGTAGTTTCC
Slc2a4 U	CATGGCTGTCGCTGGTTTC	Slc2a4 L	AAACCCATGCCGACAATGA
Soat1 U*	CTGGGGAGAATCCTGAGCAAG	Soat1 L	GAGTGCACACCCACCATTGTC
Srebf1 U	ACAGCGGTTTTGAACGACATC	Srebf1 L	GGAAGTCACTGTCTTGGTTGTTGA
TNFa U	CACAAGATGCTGGGACAGTGA	TNFa L*	GGCTCTGTGAGGAAGGCTGT

* designed using Primer Select (DNASTAR)

2.1.4. Antibodies

Primary and secondary antibodies were used to detect Arp3 and α-tubulin in Western blot analysis of adult organs and to detect Arp3 by immunohistochemistry in preimplantation embryos.

2.1.4.1. Primary antibodies used in analysis of Arp3 mouse line

For Western blotting:

- Rabbit anti-Arp3 polyclonal antibody. Western blotting 1.5 µg/ml. Antibody was kindly provided by Theresia Stradal (GBF, Braunschweig).
- Mouse anti-α-tubulin monoclonal antibody (#T9026 DM 1A, Sigma). Western blotting 1:5000.

For immunohistochemistry:

- Mouse anti-Arp3 IgM monoclonal antibody (49B6). Epitopes: amino acids 55-66 and 82-93. Immunohistochemistry 1:8. Antibody was kindly provided by Theresia Stradal (GBF, Braunschweig).
- Mouse anti-Arp3 IgG monoclonal antibody (#612134, BD Biosciences). Immunogen: amino acids 181-291. Immunohistochemistry 1:100.

2.1.4.2. Secondary antibodies used in analysis of Arp3 mouse line

For Western blotting:

- Goat anti-Rabbit IgG coupled to HRP (#A0545, Sigma). Epitope: amino acids 351-365. Western blotting 1:5000.
- Goat anti-Mouse IgG coupled to HRP (#A9044, Sigma). Western blotting 1:5000.

For immunohistochemistry:

- Horse anti-Mouse IgG (H+L) coupled to HRP (#PI-2000, Vector Laboratories). Immunohistochemistry 1:200.

2.1.5. Databases

Ensembl	www.ensembl.org
Expert Protein Analysis System	www.expasy.ch/sprot/
International Federation of Pigment Cell Societies	ifpcs.med.umn.edu/micemut
International Gene Trap Consortium	www.genetrap.org
NCBI Databases	www.ncbi.nlm.nih.gov

2.1.6. Gene cDNA sequences

Arp3: GenBank accession number NM 023735

KRAB-ZF: IMAGE clone 894250

Lsm16: GenBank accession number NM 153799

Tyrosinase: GenBank accession number NM 011661

2.1.7. Mouse strains

Three mouse strains, C57BL/6J, 129/SvS2 Hsd, and CD1 were used to perpetuate the gene trap mutations in the mouse lines analyzed in this thesis. Mice were kept at 21-23°C, under standard conditions on a 5 a.m. – 7 p.m. light cycle, 7 p.m. - 5 a.m. dark cycle and were supplied with water and feed ad libitum.

2.2. Methods

2.2.1. Molecular biology

2.2.1.1. Isolation of genomic DNA

Genomic DNA was isolated from mouse tissues for the purpose of genotyping. Where larger amounts on DNA were needed for genotyping by Southern blotting (i.e. KRAB-ZF line), the phenol/chloroform/isoamyl-alcohol extraction was used. This allowed for utilization of larger tissue pieces. Although the DNA extracted by this method also proved to be a good template for PCR genotyping, the preferred method for obtaining DNA for PCR genotyping was the Hot Shot extraction.

2.2.1.1.1. Phenol/chloroform/isoamyl-alcohol extraction

Genomic DNA was isolated from tail biopsies and yolk sacs for genotyping mice and embryos, respectively, of the KRAB-ZF line. The starting material was a 1 cm piece of mouse tail or a yolk sac that had been centrifuged and the supernatant removed. The tail or yolk sac was digested overnight at 55°C in DNA lysis buffer (100 mM Tris-HCl pH 8.5, 200 mM NaCl, 0.2% SDS, 5 mM EDTA, proteinase K 0.5 mg/ml). The DNA was extracted using phenol/chloroform/isoamyl-alcohol (25:24:1) and precipitated using 100% ethanol. Following washing the pellet in 70% ethanol and drying, the DNA was resuspended in TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). This DNA was used for Southern blot genotyping the KRAB-ZF line.

2.2.1.1.2. Hot Shot extraction

For PCR genotyping, only small amounts of genomic DNA are necessary, or in the case of blastocysts, only small amounts of DNA are present. For this, genomic DNA was isolated from 1 mm tail biopsies and blastocysts by Hot Shot extraction as described by Truett *et al*. (2000). For tails, 200 µl 50 mM NaOH and 50 µl 1 M Tris-HCl pH 8.0 were used, whereas for blastocysts, amounts were reduced to 1/10th. This DNA was used for PCR genotyping each of the mouse lines.

2.2.1.2. DNA extraction from agarose gels

DNA was recovered from agarose gel slices with the QIAEX II gel extraction kit (Qiagen) according to the manufacturer's protocol.

2.2.1.3. Isolation of total RNA

Total RNA was isolated from adult tissues of the *Arp3* mutant mouse line and from embryos, prenatal and adult tissues of the KRAB-ZF line by the acid-guanidinium thiocyanate-phenol-chloroform method, as described by Chomczynski and Sacchi (1987). Modifications included: pulverization of the samples in liquid nitrogen with mortar and pestle before homogenization, use of chloroform-isoamyl alcohol (24:1) instead of chloroform for the extractions, and use of DEPC-treated water for dissolving the RNA pellet.

In the *Lsm16* mutant mouse line, total RNA was isolated from adult tissues using peqGOLD RNAPure^{TM} (Peqlab) as per manufacturer's instructions.

Subsequent to both methods, the RNA concentration was determined photometrically and the integrity was confirmed by electrophoresis on a 1% denaturing agarose gel. The RNA was applied to RT-PCR and Northern blot analyses.

2.2.1.4. Quantification of nucleic acids

DNA and RNA both have their maximum absorbance at a wavelength of 260 nm. At this wavelength, an optical density (OD260) of 1 corresponds to a double-stranded

DNA concentration of 50 µg/ml, and a single-stranded DNA or RNA concentration of 38 µg/ml. Taking into consideration the dilution factor (d), the concentration of double-stranded DNA (X) was calculated using the following formula:

$$\frac{\text{OD260 x d x 50 µg/ml}}{\text{1000 µl/ml}} = \text{X µg/µl}$$

The concentration of RNA was calculated with the formula above, except the factor 38 µg/ml was used instead of 50 µg/ml.

2.2.1.5. Reverse transcription-polymerase chain reaction (RT-PCR)

RT-PCR was performed to determine differences in transcript levels of marker genes between wild-type and mutant tissues in the KRAB-ZF and *Lsm16* mutant mouse lines, as well as to detect endogenous transcripts in wild-type embryos and tissues in the KRAB-ZF line. In the KRAB-ZF line, crosses between pigmented wild-type mice were used to obtain wild-type embryos and postnatal mice. Crosses between white homozygous mutant mice were used to obtain mutant embryos and postnatal mice. Total RNA was isolated from 6-9 embryos per genotype and embryonic stage, and 3 skins per genotype and postnatal stage. In the *Lsm16* line, mice were used as mentioned in chapter 3.3.7. For RT-PCR, cDNAs were first synthesized from total RNA (2 µg) using SuperScript III reverse transcriptase (Invitrogen) and oligo$(dT)_{16}$ primers according to the manufacturer's instructions. The final volume was increased to 100 µl with sterile bidest. water (ddH_2O). The cDNAs were used as templates for PCR amplification.

In the KRAB-ZF gene trap line, endogenous KRAB-ZF transcripts were detected using primers KRAB-ZF U1 and KRAB-ZF L1 (Tab. 1). Transcripts from pigmentation marker genes were detected using the primers found in table 2. PCR reactions and conditions were as follows:

1 reaction		PCR conditions		
2.0 µl	10x PCR-buffer	15 minutes	95°C	initial denaturing
1.2 µl	25 mM $MgCl_2$	30 seconds	94°C	denaturing
0.8 µl	10 mM dNTP mix	30 seconds	57°C	annealing
1.0 µl	20 pmol gene specific primers	1-2 minutes	72°C	extension
0.2 µl	20 pmol *Rpl7* primers	10 minutes	72°C	final extension
0.3 U	Hot Star taq polymerase	hold	6°C	store
2-5 µl	cDNA	denaturing, annealing and extension were repeated for 35-40 cycles		
fill to 20 µl	sterile ddH_2O			

In the *Lsm16* gene trap line, transcripts from adipocyte and energy metabolism marker genes were detected using the primers found in table 3. PCR reactions and conditions were as follows:

1 reaction		PCR conditions		
3.0 µl	10x PCR-buffer	15 minutes	95°C	initial denaturing
1.0 µl	25 mM $MgCl_2$	30 seconds	94°C	denaturing
0.8 µl	10 mM dNTP mix	30 seconds	58°C	annealing
0.6 µl	20 pmol gene specific primers	40 seconds	72°C	extension
0.2 µl	20 pmol *Rpl7* primers	10 minutes	72°C	final extension
0.2 U	Hot Star taq polymerase	hold	6°C	store
2.0 µl	cDNA	denaturing, annealing and extension were repeated for 25-40 cycles		
fill to 30 µl	sterile ddH_2O			

Subsequent to PCR, the products were subjected to gel electrophoresis on ethidium bromide stained gels. In the case of the adipocyte and energy metabolism marker genes, the calculation of differences in transcript quantity between wild-type and homozygous mutant mice was performed by densitometric analysis.

2.2.1.5.1. Densitometric analysis of transcript quantities

Densitometric analysis was performed to quantify the differences in transcript levels of marker genes between tissues of wild-type and homozygous mutant mice. For this, PCR products were separated by electrophoresis on ethidium bromide stained gels. The gels were then photographed at 3-6 different exposure settings. From the photographs, the PCR bands were quantified using EASY Win32 software (Herolab). Here, each pair of wild-type and mutant bands was quantified together. The band of higher intensity was set at 100%; the band of lower intensity was given as a percent of the higher intensity band. The quantities were then calculated so that the wild-type band was 100% and the mutant band a percent thereof. This was conducted for the specific gene and Rpl7 (loading control) bands. The mutant value from the specific gene was divided by the mutant value from Rpl7 to determine the percent of mutant transcript relative to wild-type. The average from the measurements at 3-6 different exposure settings was calculated. This was then repeated for a second PCR reaction. The results of the analysis from both PCR reactions were then averaged to obtain the final percent of transcript quantity from mutants relative to wild-type (wild-type being 100%).

For densitometric analysis of the difference in protein levels of Arp3 between tissues of wild-type and heterozygous mice, the process was conducted as above with the exceptions that the protein was detected by Western blotting and tubulin was used as the loading control.

2.2.1.6. Northern blot

Northern blotting represents a further method, in addition to RT-PCR, for verifying gene expression. Results shown in this thesis represent Northern blotting conducted on tissues of the *Arp3* and *Lsm16* gene trap lines. For this, 20 µg of total RNA was loaded per lane on a 1% agarose gel with 2% formaldehyde, electrophoresed in Running buffer (20 mM MOPS, 5 mM sodium acetate pH 5.3, 5 mM EDTA, 7.4% formaldehyde), vacuum blotted in 10× SSC onto a positively charged Hybond-XL nylon membrane (Amersham), and UV cross-linked. Hybridization was performed using Rapid-hyb buffer (Amersham) with a gene specific probe radiolabeled with α^{32}P-dCTP using the Megaprime DNA labeling kit (Amersham) per manufacturer's

instructions. Radioactive signals were detected on Hyperfilm (Amersham). The 28S rRNA band was used as a loading control.

The probe used to detect the *Arp3* transcript consisted of a Bgl II – Pst I fragment of 1239 bp. This corresponds to positions 1002-2240 of the *Arp3* cDNA sequence, NCBI GenBank accession: NM 023735. This is 3´ to the vector integration site and includes part of exon 8, all of exons 9-11 and most of exon 12.

The probe used to detect the *Lsm16* transcript consisted of an Acc I – Ava I fragment of 643 bp. This corresponds to positions 208-850 of the *Lsm16* cDNA sequence, NCBI GenBank accession: NM 153799. This is 5´ to the vector integration site and includes part of exon 2, all of exon 3 and part of exon 4.

2.2.1.7. Sequencing of KRAB-ZF and tyrosinase transcripts

Sequencing reactions were prepared using the Big Dye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). Separation of PCR products was performed with an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) and sequence analysis was conducted with ABI PRISM Sequence Analysis v3.0 software (Applied Biosystems). The primers used for sequencing are listed in table 1.

2.2.2. Genotyping protocols

Genotyping of mice and embryos of the *Arp3*, KRAB-ZF and *Lsm16* mouse lines was based on the standard PCR genotyping protocol as follows:

1 reaction	
2.5 µl	10x PCR-buffer
1.0 µl	25 mM $MgCl_2$
0.8 µl	10 mM dNTP mix
1.0 µl	20 pmol gene specific primers
1.0 µl	20 pmol vector primer
0.2-0.3 U	r taq polymerase
2.0 µl	DNA
fill to 25 µl	sterile ddH_2O

PCR conditions		
5 minutes	94°C	initial denaturing
30 seconds	94°C	denaturing
30 seconds	59°C	annealing
1 minute	72°C	extension
5 minutes	72°C	final extension
hold	6°C	store
denaturing, annealing and extension were repeated for 36 cycles		

Mouse line specific changes to this protocol are listed below.

2.2.2.1. *Arp3* mice and preimplantation embryos

Mice were genotyped at the *Arp3* locus by PCR using the standard genotyping protocol with the primers Arp3 U1, Arp3 L1 and LacZ L2 at an annealing temperature of 53°C.

For preimplantation embryos, PCR was conducted as for mice, with the exceptions that the component amounts in the PCR reaction were doubled to a final volume of 50 µl and 8 µl DNA were used. PCR was performed for 40 cycles.

2.2.2.2. KRAB-ZF mice and embryos

Mice and embryos from crosses between wild-type and heterozygous mice were genotyped by PCR to detect the *lacZ* gene from the gene trap vector. PCR was conducted using the standard protocol, but only the *lacZ* primers LacZ U1 and LacZ L1 were used.

Mice and embryos from heterozygous matings were genotyped at the trapped KRAB-ZF locus using Southern blot analysis. For this, genomic DNA was digested using the restriction endonuclease Bam HI, loaded on a 0.7% agarose gel and electrophoresed in TBE buffer (0.9 M Tris base, 0.9 M boric acid, EDTA pH 8.0). The DNA was then vacuum blotted onto a Hybond-XL nylon membrane (Amersham) as follows: 10-20 minutes in Depurination solution (0.2 N HCl), 30-40 minutes in Denaturation solution (0.5 N NaOH, 0.5 M NaCl), 30 minutes in Neutralization solution (0.5 M Tris pH 7.4, 1.5 M NaCl) and finally 30-60 minutes in Transfer solution (20× SSC). Following UV cross-linking, the membrane was hybridized using Rapid-hyb buffer (Amersham) with a KRAB-ZF gene specific probe radiolabeled with α^{32}P-dCTP using the Megaprime DNA labeling kit (Amersham) per manufacturer's instructions. Radioactive signals were detected on Hyperfilm (Amersham).

The probe used to detect the KRAB-ZF transcript consisted of a Sal I – Bgl II fragment of 502 bp. This corresponds to positions 1-501 of the KRAB-ZF cDNA sequence of the IMAGE clone 894250. This is 5´ to the vector integration site and includes part of exon 2, all of exon 3 and part of exon 4.

Mice of the KRAB-ZF gene trap line were genotyped at the *p* locus by PCR. Reactions and conditions were those of the standard protocol, except the *p* gene-specific primers p U2 and p L2 were used, and PCR was conducted for 40 cycles.

2.2.2.3. *Lsm16* mice

Mice were genotyped at the *Lsm16* locus by PCR using the standard protocol. The *Lsm16* gene-specific primers Lsm16 U1 and Lsm16 L1, as well as the *engrailed2* intron-specific primer En2 in L1 were used.

2.2.3. Cell culture

2.2.3.1. Blastocyst cultivation

To study the morphology and development of blastocysts up to lethality in the *Arp3* mutant mouse line, embryos from heterozygous matings were isolated at day 3.5 postcoitum in M2 medium (Sigma), divided into groups based on developmental variation, photographed and cultivated overnight in M16 medium (Sigma) at 37°C and 10% CO_2. The next day, each group was divided into subgroups based on whether development had progressed or not. The embryos were photographed and genotyped by PCR.

2.2.4. Histology

2.2.4.1. Paraffin sections of embryos and adult tissues

Paraffin sectioning was used to study the pigmentation of eyes and skin (hairs), as well as the expression of the β-galactosidase fusion protein in stained embryos in the A20010 line and muscle morphology in the A008A01 line. The samples were first fixed. Eyes, skin and stained embryos were fixed in 4% PFA overnight. Muscle was fixed 2 weeks in Bouin's fixative. Fixed specimens were dehydrated through a graded series of ethanol, cleared in xylol, then infiltrated and embedded in paraffin. Serial sections were cut and dried for 4 hours at 37°C or overnight at room temperature and stored at 5°C until processing. Embryo sections were counter-stained in eosin, whereas adult eye sections were stained with hematoxylin and eosin.

2.2.4.2. Cryostat sections of adult tissues

Cryosectioning was used for studying the expression of the β-galactosidase fusion protein and mATPase activity in mouse tissues. Tissues were dissected from the mouse, attached to a cork disk using Polyfreeze tissue freezing medium (Polysciences), snap-frozen in liquid nitrogen cooled isopentane and mounted to the cryostat chuck. Serial sections were cut at -20°C for muscle and -30°C for adipose tissues, dried for 2 hours at room temperature, then stored at -80°C or immediately stained for β-galactosidase or mATPase activity.

2.2.4.3. Hematoxylin and Eosin staining of paraffin sections

Paraffin sections were deparaffinized in Roticlear (Roth), then rehydrated in decreasing concentrations of ethanol i.e. 96%, 80%, 50%, 30%. Each of these steps was performed 2 times for 2 minutes. Following 5 minutes in tap water, the sections were stained for 8-10 minutes in Mayer's hematoxylin (Sigma), washed shortly in tap water, differentiated in acid ethanol (1% HCl in 70% EtOH), then washed 10 minutes in running tap water. The sections were counter-stained 30-60 seconds in Eosin Y (1% in 0.05% acetic acid) (Sigma), differentiated for 30 seconds in 70% ethanol, then dehydrated in 80% and 96% ethanol, followed by Roticlear. These steps were again 2 times for 2 minutes. The sections were mounted in entellan (Merck), analyzed and photographed.

2.2.4.4. X-gal staining to detect β-galactosidase expression in tissue sections

To detect the fusion protein in tissue sections, cryostat sections were fixed in Staining fixative (1x PBS, 0.01% DOC, 0.02% NP40, 2 mM $MgCl_2$, 0.2% glutaraldehyde, 1% formaldehyde) for 10 minutes at 4°C, washed in Staining wash (1x PBS, 0.01% DOC, 0.02% NP40, 2 mM $MgCl_2$) 3 times for 10 minutes (2 times at 4°C and 1 time at room temperature), then stained overnight in X-gal solution (1x PBS, 0.01% DOC, 0.02% NP40, 2 mM $MgCl_2$, 20 mM Tris pH 8.0, 5 mM $K_4Fe(CN)_6$ $3H_2O$, 5 mM $K_3Fe(CN)_6$, 1 mg/ml X-gal in DMF) at 30°C to reveal the β-galactosidase activity. Following staining, sections were washed in PBS and ddH_2O, counter-stained 3-5 minutes with Nuclear Fast Red (0.1% in 5% aluminum sulfate), dehydrated in a series of increasing ethanol concentrations i.e. 30%, 50%, 70%, 90%, 96% and once in isopropanol, each 1 minute, then 10 seconds in xylol and mounted using PolyMount (Polysciences). After overnight drying, the sections were analyzed and photographed.

2.2.4.5. Bluo-gal staining to detection β-galactosidase expression in embryos

Following matings, the day on which the vaginal plug was detected was considered to be day 0.5 of gestation (E0.5). Embryos were isolated into sterile ice-cold phosphate-buffered saline solution (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO), then stained similar to the protocol for X-gal staining of tissue sections. Embryos were fixed in Staining fixative for 1-3 hours at 4°C for embryos up to stage E12.5. Older embryos were fixed overnight, then cut in half and refixed 1 hour. All embryos were washed in Staining wash 2 times for 20 minutes at 4°C and 1 time for 20 minutes at room temperature, then stained overnight in Bluo-gal solution (1x PBS, 0.01% DOC, 0.02% NP40, 2 mM $MgCl_2$, 20 mM Tris pH 8.0, 5 mM $K_4Fe(CN)_6$ $3H_2O$, 5 mM $K_3Fe(CN)_6$, 1 mg/ml Bluo-gal in DMF) at 30°C to reveal the β-galactosidase activity. The embryos were then washed in PBS and photographed.

To better observe the internal staining, embryos were either paraffin sectioned or cleared. Clearing consisted of dehydrating the embryos in a series of increasing ethanol concentrations for 4-5 hours each 30%, 50%, 70% and 80%, then overnight in 90% and 100%. Dehydrated embryos were then transferred to a mixture of benzyl-benzoate and benzyl-alcohol (2:1). After several hours of clearing, depending on size, the embryos were ready for photographing.

2.2.4.6. Myofibrillar ATPase staining

To differentiate between fiber types in skeletal muscle, myofibrillar ATPase staining was performed according to Ogilvie and Feeback (1990). This consisted of cryostat sections from muscle being treated with Preincubation medium No. 1 (pH 4.5) for 6 minutes, washed in Tris buffer three times for 2 minutes, then treated with Incubation medium No. 2 (pH 9.4) for 25 minutes. Next, the sections were washed 10 minutes in 1% calcium chloride solution, stained in 0.1% toluidine blue for 40 seconds, rinsed in distilled water, dehydrated in 95% ethanol, then 100% ethanol, cleared in xylene and mounted in entellan (Merck). Differences in ATPase activity among fibers leads to differences in staining intensities of toluidine blue (Sigma). This allows for the identification of type I and type II fibers.

2.2.4.7. Photography

Sections were viewed with a Leica DMRBE upright microscope and photographed using a ProgResC12 camera and software (Jenoptik). Embryos were photographed using a Leica MZFLIII stereomicroscope with Polaroid PDMC-3 digital camera and DMC-2 software.

2.2.5. Biochemistry

2.2.5.1. Protein extraction from tissues

The tissues to be analyzed were isolated from the mouse, weighed and washed in a cold solution of PBS and proteinase inhibitors (pepstain A, AEBSF, trasylol, leupeptin, each 1:1000). The tissues were then cut into small pieces and placed in liquid nitrogen. Following pulverization in liquid nitrogen with mortar and pestle, the tissues were weighed and resuspended in 2 ml RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% triton X-100, 0.025% DOC, 1 mM EGTA) per 1 g tissue. These were then homogenized by repeatedly passing through a syringe. The homogenates were cleared by centrifuging at 4°C for 10 minutes at 13000 x g. The supernatant was retained and the protein concentration was determined according to Bradford, (1976). The protein extracts were analyzed by SDS-PAGE and subsequent Western blotting.

2.2.5.2. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

For the separation of proteins by molecular weight, SDS-PAGE was performed essentially as described previously (Laemmli, 1970). For this, the protein extracts were adjusted to a concentration of 4 µg/µl using SDS sample buffer (375 mM Tris-HCl pH 6.8, 2% SDS, 12% glycerin, 0.05% bromphenol blue, 10% β-mercaptoethanol), heated ~5 minutes at 95°C, then loaded to the gel. The “High molecular weight marker” (Sigma) was used to determine sample molecular weights. Running gels of 10% acrylamide solution were used. Electrophoresis was conducted at 20-30 mA until the bromphenol blue band reached the bottom. Following separation, the proteins were subjected to Western blot analysis.

2.2.5.3. Western blot

For comparing protein expression level differences between organs of wild-type and heterozygous mice, Western blot analysis was conducted. Following SDS-PAGE, the separated proteins were transferred to a nitrocellulose membrane (Pall Corporation). This involved equilibrating the membrane and gel in Transfer buffer (25 mM Tris-HCl pH 8.5, 150 mM glycine, 10% methanol) and blotting 1 hour at 100 mA using a semi-dry blotting apparatus (Biometra). After blotting, membranes were stained in Ponceau S (Sigma) to determine the efficiency of the transfer and to allow for marking of the molecular weight marker.

To detect the proteins of interest using antibodies, the blotted membranes were first blocked for 1 hour at room temperature or overnight at 4°C in Blocking buffer (5% powdered milk in TBST buffer (20 mM Tris-HCl pH 7.6, 140 mM NaCl, 0.1% Tween 20)). This was to prevent the nonspecific adsorption of antibodies to the membrane. The membranes were then rinsed in TBST buffer and incubated with the primary antibody in TBST buffer with 1% BSA for 1 hour at room temperature. Next, the membranes were washed shortly in TBST buffer, then twice for 15 minutes in TBSX buffer (20 mM Tris-HCl pH 7.6, 140 mM NaCl, 1% Triton X-100), followed by incubation for 1 hour with the secondary antibody in TBST buffer with 1% BSA. The membranes were then washed again as after the primary antibody. The horseradish peroxidase-linked secondary antibody was detected using Enhanced Chemiluminescence (ECL) with ECL Western blotting detection reagents (Amersham) as per manufacturer's instructions. Protein bands were visualized on Hyperfilm ECL (Amersham). The calculation of differences in protein quantity between organs of wild-type and heterozygous mice was performed by densitometric analysis (see 2.2.1.5.1).

2.2.5.4. Immunohistochemistry

For the *in vivo* visualization of proteins in preimplantation embryos of the *Arp3* mutant mouse line, immunohistochemistry was conducted. This involved the isolation of embryos at day 3.5 postcoitum from females of heterozygous matings. The embryos were isolated in M2 medium (Sigma), photographed and the zona pellucida was removed using Tyrode's solution (Sigma). The embryos were then fixed in 4% paraformaldehyde in PBS for 20 minutes and permeabilized for 20 minutes with 0.1% Triton X-100 in PBS. After incubation for 30 minutes in Blocking buffer (0.1% Triton

X-100, 1% BSA in PBS), embryos were incubated for 1 hour with the primary antibody in PBS with 1% BSA. The embryos were then washed 3 times for 5 minutes in Blocking buffer and then incubated for 1 hour with the secondary antibody in PBS with 1% BSA. Next, the embryos were washed 5 minutes in Blocking buffer, then 0.1% Triton X-100 in PBS, then PBS and finally 0.1 M Tris pH 7.2. Visualization of the horseradish peroxidase-linked secondary antibody was performed using DAB (Sigma). Embryos were viewed with a Leica DMRBE upright microscope and photographed using a ProgResC12 camera and software (Jenoptik). The embryos were genotyped by PCR.

3. Results

3.1. Analysis of the *Arp3* gene trap mouse line A009F03

Although research to characterize Arp3 and especially the Arp2/3 complex has been underway worldwide for more than 10 years, in the mouse the analysis of Arp3 on the whole organism is not found in the literature. Therefore, the analysis carried out in this work is centered on the characterization of Arp3 with respect to the mouse as a model system.

3.1.1. Site of vector integration in the trapped *Arp3* gene

Work conducted by Elke Freese (Diploma thesis: Freese, 2003) shows that the mutant mouse line A009F03 has the PT1βgeo vector (Fig. 2A) integrated into intron 1 of the *Arp3* gene (Fig. 2B). The *Arp3* cDNA sequence was BLASTed to the Mouse Genome in order to obtain the intron-exon structure of the gene (Fig. 2B). The *Arp3* gene is located on chromosome 1 and consists of 12 exons spanning almost 50 kb of DNA. The mRNA is 2525 bp long with an open reading frame (ORF) of 1257 bp. This ORF codes for a protein of 418 amino acids with a calculated mass of 47.4 kDa.

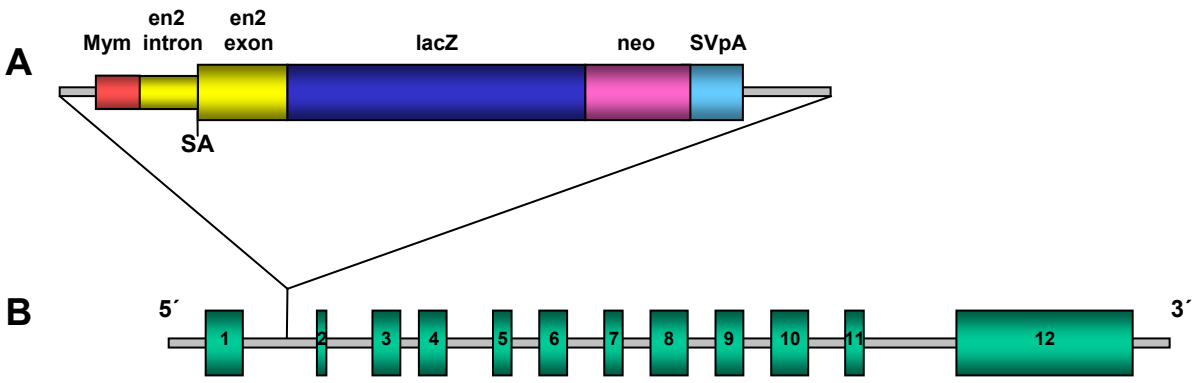


Figure 2. Schematic representation of the PT1βgeo splice-acceptor vector and the site of its integration in the *Arp3* locus. (A) The PT1βgeo vector with *βgeo* cassette. The Mym sequence (red) represents a 109 bp section of the *engrailed2* intron that has been found to be in reverse orientation due to the integration event. (B) The vector has integrated into intron 1 of the *Arp3* locus on chromosome 1. The *Arp3* locus consists of 12 exons (green). en2: *engrailed2*, SA: splice-acceptor site, lacZ: *β-galactosidase* gene, neo: *neomycin phosphotransferase* gene, SVpA: simian virus polyadenylation signal. (Adapted from Freese, 2003).

Since the gene trap mutation is recessive embryonic lethal, analysis of adult mice was conducted only on wild-type and heterozygous animals.

3.1.2. Expression of *Arp3* transcripts and protein is reduced in tissues of heterozygous mice

To characterize the distribution of the *Arp3* mRNA in tissues of adult wild-type mice, Northern blot analysis was conducted using total RNA. All organs analyzed were found to express the *Arp3* mRNA at the expected size of 2.5 kb (Fig. 3A). The highest transcript levels were observed in spleen, thymus, lung and intestine. The lowest levels were found in liver and muscle. The 28S rRNA band was used as loading control.

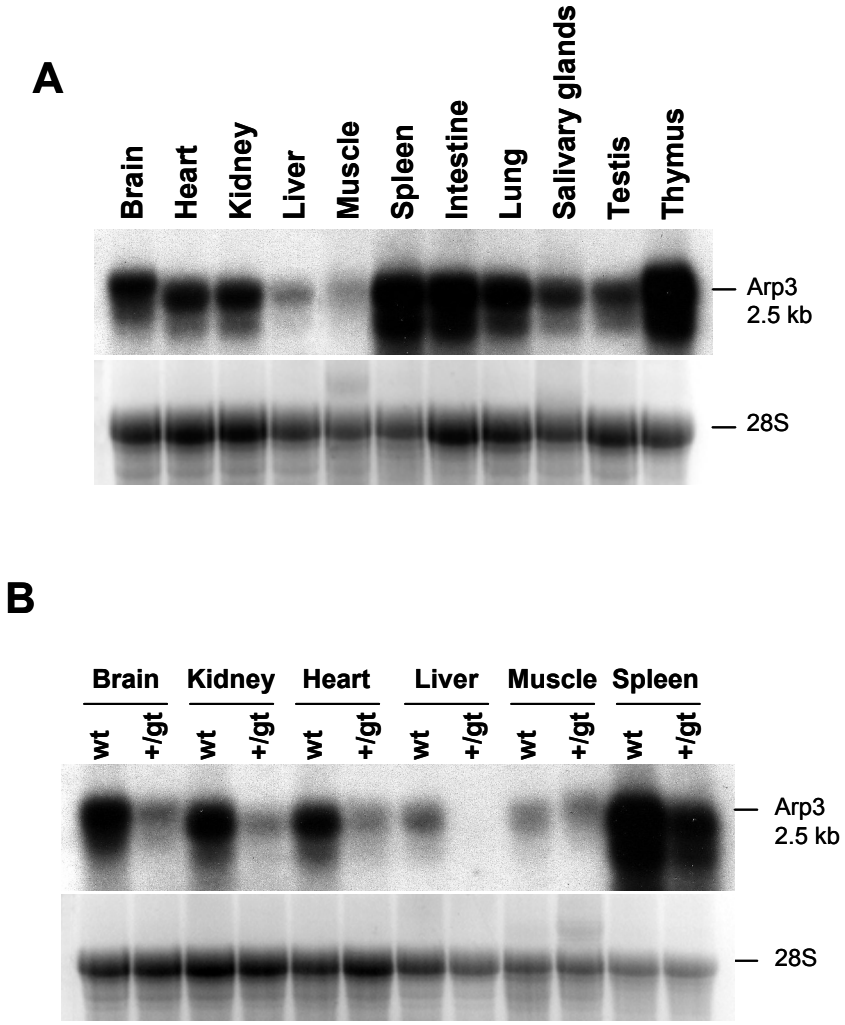


Figure 3. Northern blot analysis of *Arp3* mRNA in adult mouse tissues.
(A) Total RNA (20µg) from 11 organs of wild-type adult mice was analyzed by Northern blotting using a probe 3´ to the vector integration site. The *Arp3* transcript was found in all organs tested at the expected size of 2.5 kb. (B) Northern blot results comparing *Arp3* mRNA amounts in a select group of wild-type and heterozygous organs indicate that *Arp3* transcript quantities are reduced in heterozygotes as compared to wild-types. *Arp3* transcript in the heterozygous liver was detectable only after a long expose time. The 28S-ribosomal RNA band was used as quantity control.

To compare transcript levels between wild-type and heterozygous mice, Northern blot analysis was also conducted using total RNA from the organs: brain, kidney, heart, liver, skeletal muscle and spleen of heterozygous mice (Fig. 3B). As expected, transcript levels in heterozygous mice were reduced by approximately 50% in all organs. This provides evidence that the mutation markedly diminishes *Arp3* mRNA accumulation.

To verify that the amount of Arp3 protein was also reduced in heterozygous mice, analysis by SDS-PAGE with Western blotting was conducted using protein extracted from wild-type or heterozygous heart, liver and muscle. Arp3 of the expected size of 47.4 kDa was detected in all organs analyzed. The amount of Arp3 protein was reduced in heterozygous animals, as compared to wild-type (Fig. 4A). Similar results were observed in brain and spleen (data not shown). Arp3 protein levels were quantified by densitometric analysis and found to be reduced in heterozygous mice by 47%, 55% and 54% in heart, liver and skeletal muscle, respectively (Fig. 4B). α-Tubulin served as loading control. From the reduction of Arp3 by approximately 50% in heterozygous mice, it can be inferred that homozygous mutants display an Arp3-null phenotype. This is supported by the prediction that the Arp3/β-gal fusion protein possesses only the first 14 amino acids from the endogenous Arp3 protein.

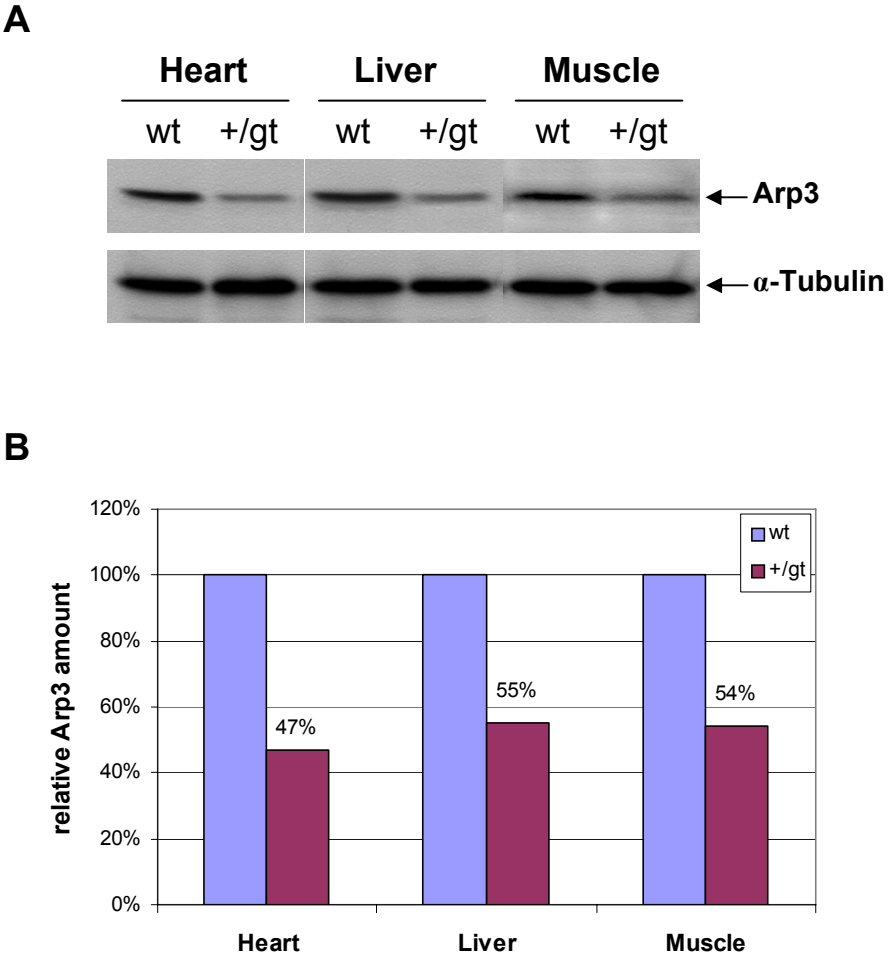


Figure 4. Quantitation of Arp3 protein in tissues of wild-type and heterozygous adult mice. (A) Protein extracts (40 µg) from heart, liver and skeletal muscle were analyzed by SDS-PAGE and subsequent Western blotting. Arp3 was detected using a rabbit polyclonal anti-Arp3 antibody (gift from T. Stradal), followed by a HRP labeled anti-rabbit secondary antibody (Sigma). Detection was performed using an ECL detection kit (Amersham). α-Tubulin was used as loading control. The highest quantity of Arp3 was found in liver, followed by heart and skeletal muscle. In all three organs, heterozygotes (+/gt) appear to have a reduced amount of Arp3 in comparison to wild-types (wt). (B) Densitometric analysis following Western blotting shows that Arp3 quantities in organs are reduced by approximately 50% in heterozygous mice as compared to wild-type. For each organ, the Arp3 amount in wild-type was set as 100%.

3.1.3. *Arp3* homozygous mutants are embryonic lethal around stage E3.5

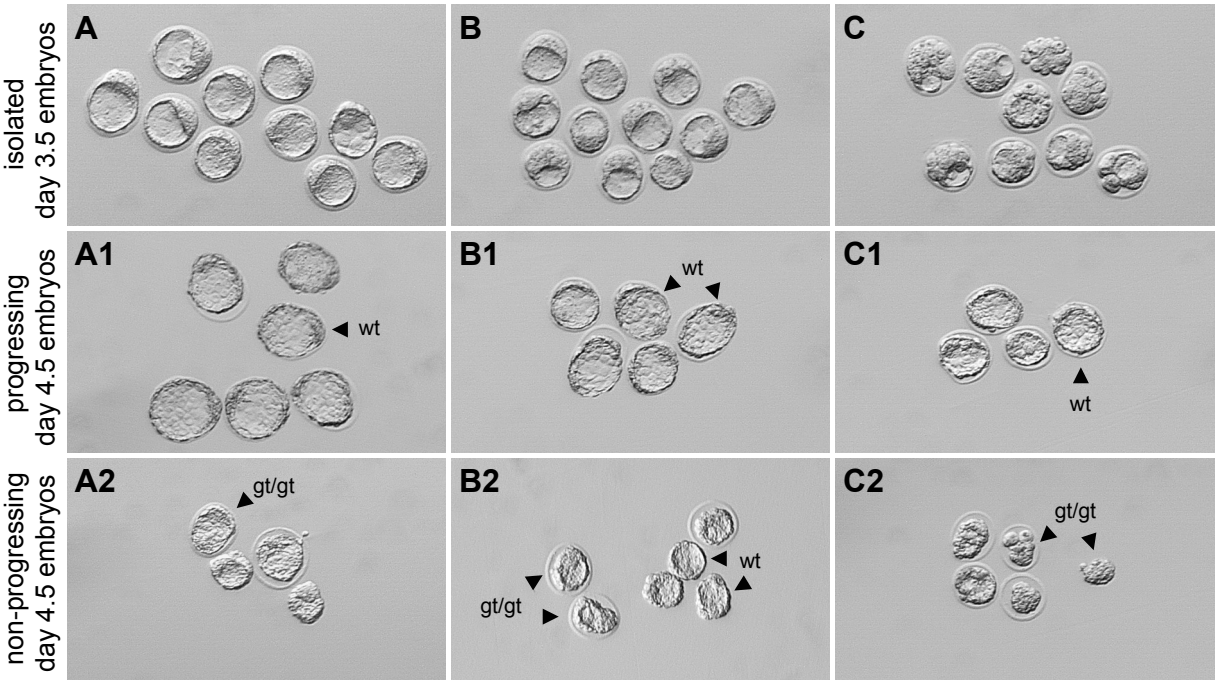
Preliminary analysis had indicated that the *Arp3* mutation causes lethality between stages E3.5 and E8.5 (Freese, 2003). Because implanted blastocysts are difficult to prepare from the uterus without substantial contamination by maternal tissue, reliable genotyping at this stage is complicated. Therefore, preimplantation embryos were cultivated *in vitro* to examine the time of early lethality more precisely.

Embryos from heterozygous matings were isolated at day 3.5 postcoitum and divided into 3 groups (A, B or C) based on variations in the embryo's development (Fig. 5A-C). The furthest developed embryos were put into group A. This group consisted primarily of almost fully expanded blastocysts. The least developed embryos were represented in group C including compact morulae and embryos beginning to develop the blastocoel. Embryos intermediate in development were collected in group B. This group consisted of expanding blastocysts. All embryos were cultivated overnight and continuing development was assessed the next day. Embryos had either progressed to later developmental stages (Fig. 5 "progressing" A1-C1), or arrested or even started to disintegrate (Fig. 5 "non-progressing" A2-C2).

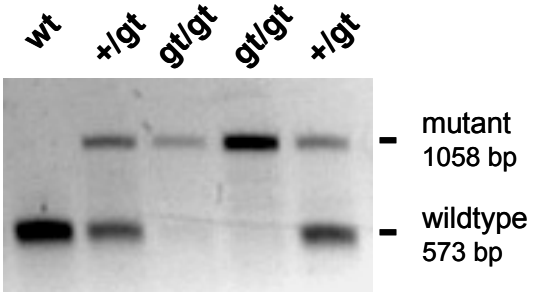
When the cultivated embryos were genotyped by PCR using gene specific primers to distinguish wild-type from mutant alleles (Fig. 5D), it was found that all homozygous mutants failed to develop further, whereas most wild-type and heterozygous embryos progressed through development normally.

This experiment was conducted 3 times totaling 85 embryos. Significantly, 49 of 71 wild-type or heterozygous embryos continued to develop *in vitro*, whereas all 14 homozygous mutant embryos did not (Fig. 5E). This result strongly indicates that Arp3 is absolutely required for early mouse development, although homozygous mutant embryos were able to develop until E3.5 with no apparent phenotype.

Comparing the morphology of embryos in each of the experimental groups A, B and C in figure 5, it is not possible to distinguish E3.5 homozygous mutants from wild-type. Surprisingly, homozygous mutants, which supposedly do not synthesize Arp3 protein, still develop fairly normal to stage E3.5. This raises the question of the requirement of Arp3 up to the formation of blastocysts.



D



E

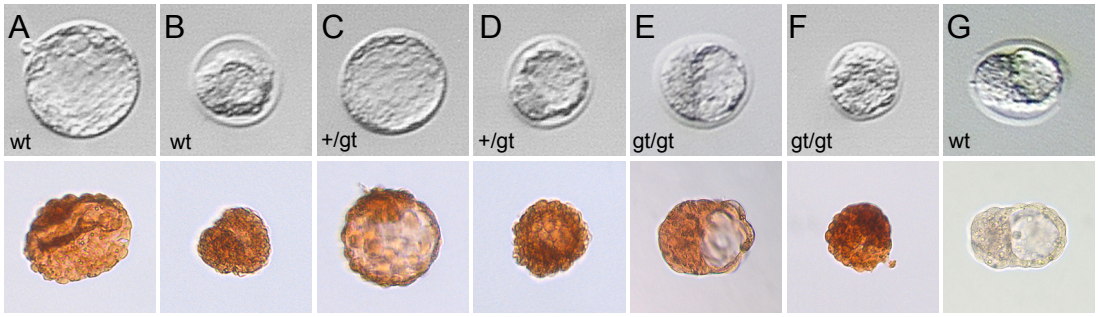
	wt	+/gt	gt/gt
progressing (49)	13	36	0
non-progressing (36)	7	15	14
total (85)	20	51	14

Figure 5. Analysis of blastocyst development in overnight culture.
Embryos from heterozygous matings were isolated at day 3.5 postcoitum and divided into 3 groups (A, B or C) based on the embryo's development. The furthest developed embryos were put into group A, the least developed into group C and those intermediate into group B. Following overnight cultivation, each group was further divided into two subgroups based on whether the embryo continued to develop ("progressing" A1-C1) or ceased development ("non-progressing" A2-C2). In the subgroup photographs, the wild-type (wt) and homozygous mutants (gt/gt) are labeled. Non-labeled blastocysts are heterozygous. (D) PCR genotyping verifies the presence of homozygous mutant embryos. The product of the wild-type allele is 573 bp and the product of the mutant allele is 1058 bp. (E) From 3 separate experiments, the distribution of 85 genotyped embryos shows that no gt/gt embryos developed further in culture. All 14 isolated mutant embryos stopped development and began deteriorating before day 4.5. More than half of the wild-type (wt) and heterozygous (+/gt) embryos developed further after isolation (progressing), whereas the remaining did not develop further (non-progressing). The distribution of the genotypes was not significantly different from the expected Mendelian distribution.

3.1.4. Arp3 protein is present in preimplantation embryos

To test for Arp3 protein in preimplantation blastocysts, embryos from heterozygous matings were isolated at day 3.5 postcoitum and analyzed by immunohistochemistry.

Photographs were taken before and after the immunohistochemical staining (Fig. 6A-G). As observed previously, slightly different developmental progression was observed among the embryos regardless of whether they were wild-type, heterozygous or homozygous mutants. Genotype distributions were close to the expected Mendelian ratio (Fig. 6H). Surprisingly, however, embryos of all genotypes contained Arp3 protein since homozygous and heterozygous mutants stained equally intense with the antibody as wild-type embryos. This observation then explains why mutant embryos develop normally to E3.5 although they carry a null-mutation in the essential Arp3 gene. The most likely explanation for the presence of Arp3 in early embryos seems that sufficient Arp3 protein and mRNA are supplied by the mother. This maternal contribution seems to rescue the early developmental stages.



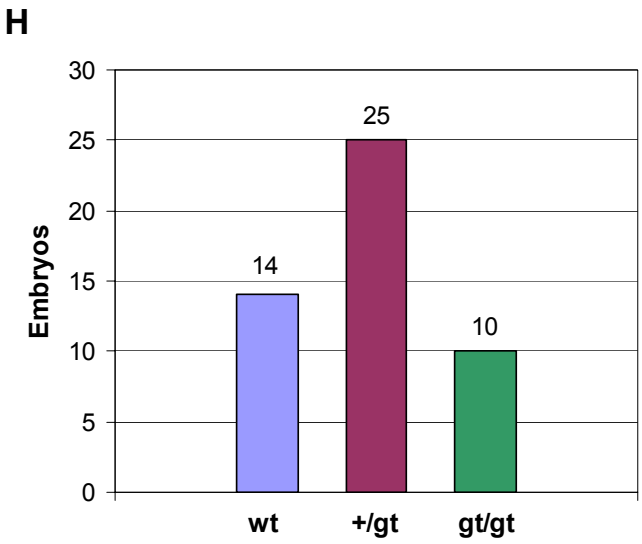


Figure 6. Immunohistochemical detection of the Arp3 protein in preimplantation embryos.
(A-F) Embryos from heterozygous matings were isolated at day 3.5 pc, photographed (upper row) and then analyzed by immunohistochemical staining (lower row) to detect the Arp3 protein. Of 49 embryos the furthest developed (A, C and E) and least developed (B, D and F) with respect to each genotype are shown. Only wild-type and heterozygous embryos had developed to the fully expanded blastocyst stage. Homozygous mutants developed no further than the expanding blastocyst stage. (G) Negative control stained without the anti-Arp3 antibody. (H) At day 3.5 pc, all genotypes were found in numbers not significantly different from the expected Mendelian distribution.

3.2. Analysis of the KRAB-ZF gene trap mouse line A20010

The work on this mouse line involved the analysis of the expression of the KRAB-ZF gene, as well as the elucidation of the hypopigmentation phenotype. The putative function of the KRAB-ZF as a transcriptional repressor did not provide clues as to its possible role in the hypopigmentation phenotype.

3.2.1. Structure of the trapped KRAB-ZF gene and protein

From the results of a preliminary analysis during my Diploma Thesis (Prochnow, 2002), it was known that the PT1ATGβgeo vector (Fig. 7A) integrated into intron 4 of the putative KRAB-ZF gene (Fig. 7B). The fusion transcript in this mouse consists of exons 1-4 of the KRAB-ZF gene, followed in-frame by the vector sequences encoding *lacZ* and *neo*. The endogenous gene is composed of 6 exons encoding an mRNA of 4043 bp. The predicted open reading frame is 2055 bp.

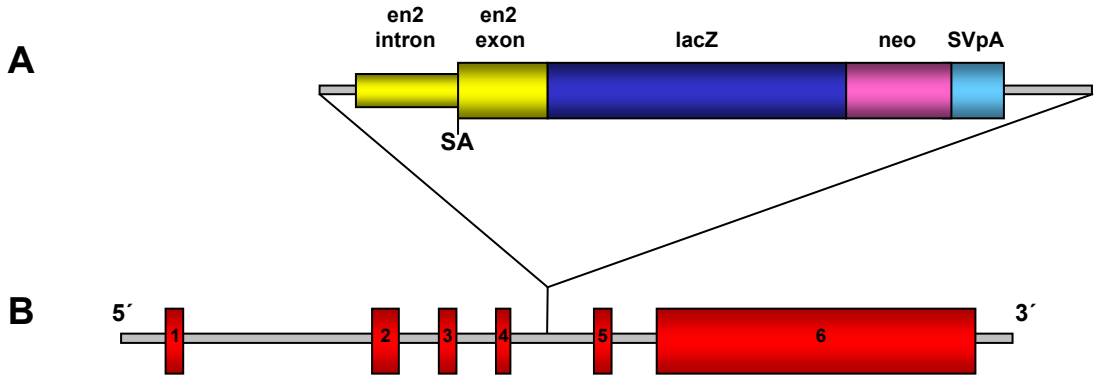


Figure 7. Schematic representation of the PT1ATGβgeo splice-acceptor vector and the site of its integration in the KRAB-ZF locus. (A) The PT1ATGβgeo vector with *βgeo* cassette. (B) The vector has integrated into intron 4 of the KRAB-ZF locus in the mouse line A20010. This locus consists of 6 exons (red). en2: *engrailed2*, SA: splice-acceptor site, lacZ: *β-galactosidase* gene, neo: *neomycin phosphotransferase* gene, SVpA: simian virus polyadenylation signal.

Since the chromosomal location of the locus was initially unknown, sequence analysis was performed on the IMAGE (Integrated Molecular Analysis of Genomes and their Expression) clone 894250, which showed 100% identity with the sequence that was trapped in A20010.

The open reading frame of the wild-type cDNA encodes a putative protein of 685 amino acids with a calculated mass of 78.7 kDa. The N-terminus of the protein has a well-conserved krüppel-associated box (KRAB) domain that consists of KRAB A and B boxes, whereas the C-terminus contains a krüppel-type zinc-finger domain with ten C_2H_2 zinc-finger motifs (Fig. 8). Each motif has a typical consensus sequence of **C**X_2**C**X_3FX_5LX_2**H**X_3**H**, with the exception of the fourth motif, which has a phenylalanine (F) substituted for a leucine (L) at position 14. A nuclear localization signal is found in the fifth zinc-finger motif. No homologies of sequences outside of the KRAB and zinc-finger domains have been found.

The site of mutation displayed in figure 8 can be deduced from the location of vector integration. The mutant fusion protein consists of the KRAB domain with both A and B boxes followed by β-galactosidase and neomycin phosphotransferase sequences derived from the vector. Since the KRAB domain remains intact, it is possible that the KRAB/β-galactosidase fusion protein can still interact with other proteins. The loss of all zinc-fingers including the nuclear localization signal suggests that the protein, however, can no longer bind DNA and is not transported into the nucleus. Thus, in contrast to wild-type KRAB-ZF proteins, the fusion protein is very unlikely to function as a transcriptional repressor.

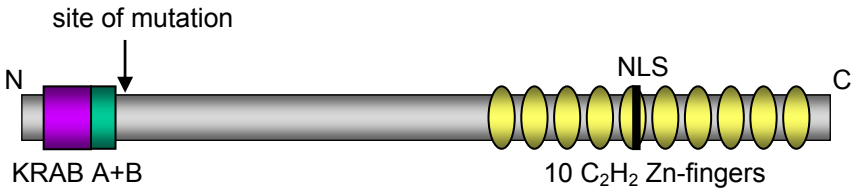


Figure 8. Schematic representation of the domains and motifs of the KRAB-ZF wild-type protein. The KRAB domain, consisting of KRAB A (purple) and B (green) boxes, is located at the N-terminus of the wild-type protein. 10 krüppel-type C_2H_2 zinc-finger motifs (yellow) are found at the C-terminus. A nuclear localization signal is found in the fifth zinc-finger (black). The arrow indicates the site of mutation, which corresponds to the location of β-galactosidase and neomycin phosphotransferase sequences in the mutant fusion protein. The section of the wild-type protein C-terminal to the site of mutation is absent in the mutant protein.

3.2.2. The KRAB/β-galactosidase fusion protein is expressed in a subset of cells in mouse embryos

To analyze the spatio-temporal expression pattern of the KRAB/β-galactosidase fusion protein, embryos in the C57BL/6J genetic background were stained for β-gal activity using Bluo-gal (ICN Biomedicals). Bluo-gal is comparable to the more commonly used X-gal, but leads to a more intense staining.

Expression of the fusion protein was found throughout the embryo in almost all tissues, but not every cell within these tissues was stained. The heart was the only organ without staining. Regions of stronger staining included the optic and otic vesicles, branchial arches, neural tube, limb buds, ventral trunk, pons and medulla oblongata. This expression pattern was seen in homozygous mutant embryos at stage E9.5 (Fig. 9A), E10.5 (Fig. 9B) and E11.5 (data not shown).

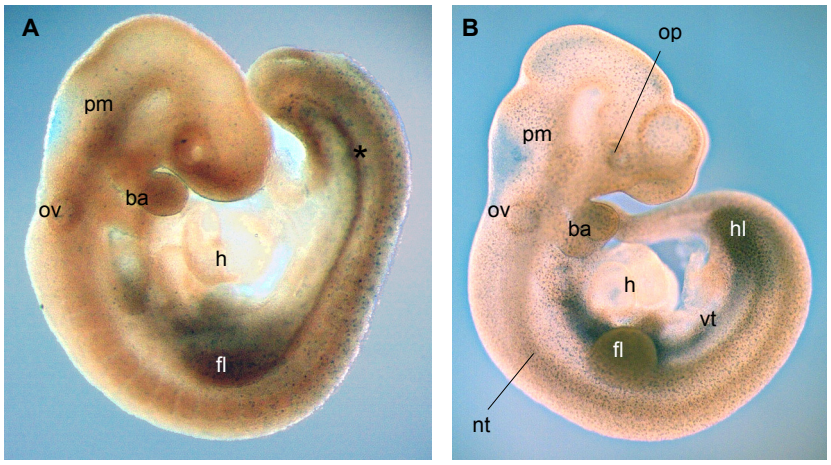


Figure 9. Expression of the KRAB/β-galactosidase fusion protein in stage E9.5 and E10.5 homozygous mutant embryos. At stages E9.5 (A) and E10.5 (B) expression is found throughout the embryo, but not in the heart (h) or in every cell. Areas of increased staining intensity are seen in the branchial arches (ba), forelimbs (fl), hindlimbs (hl), neural tube (nt), optic vesicles (op), otic vesicles (ov), pons and medulla oblongata (pm), and ventral trunk (vt). The (*) represents the site of the future hindlimb in A.

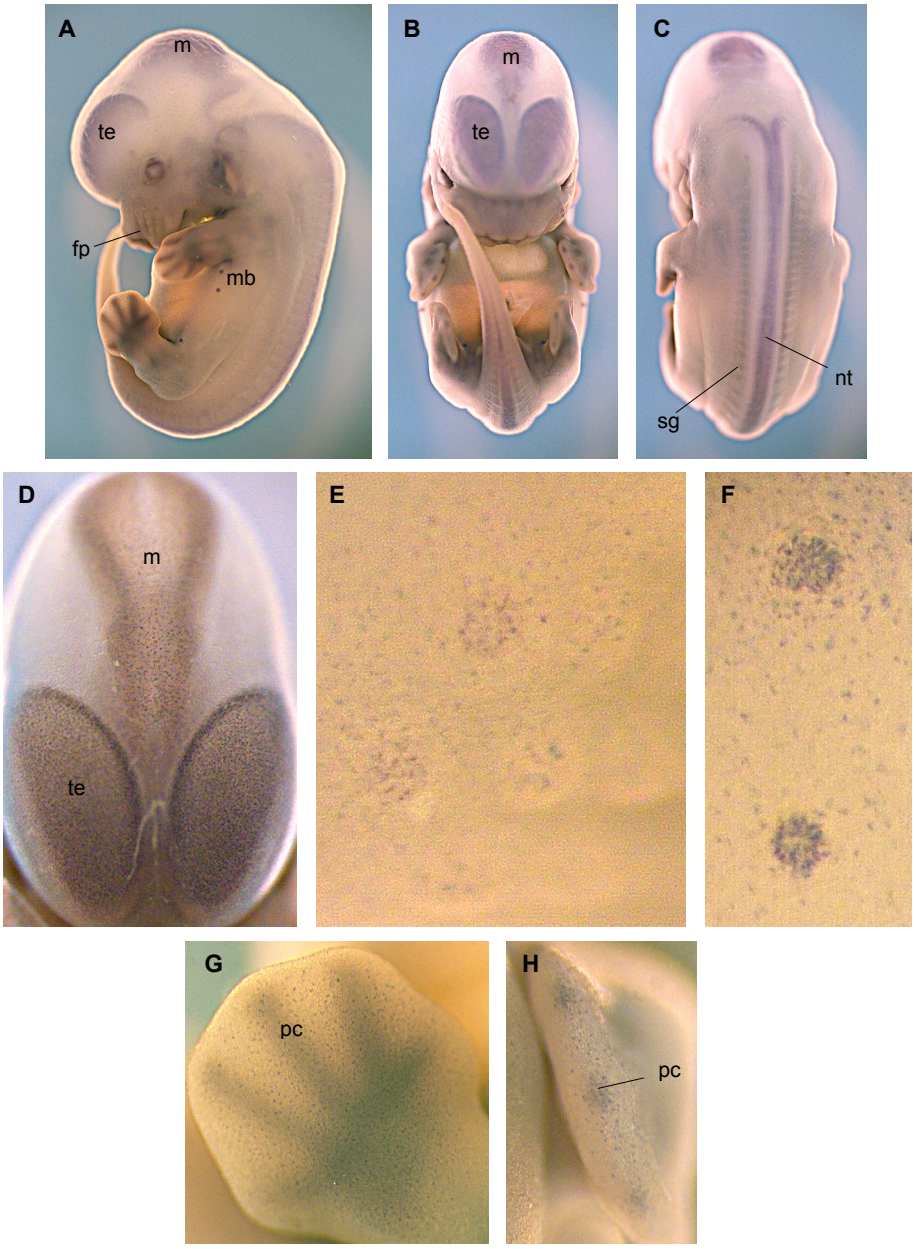


Figure 10. Expression of the KRAB/β-galactosidase fusion protein in a stage E12.5 homozygous mutant embryo. At stage E12.5, staining is seen as in earlier stages with new areas of expression in the mesencephalon (m), telencephalon (te), follicle plugs (fp), mammary buds (mb) and spinal ganglia (sg). Embryo views: lateral (A), ventral (B), dorsal (C). Higher magnification: head, dorsal view (D), follicle plugs (E), mammary buds (F), and limb buds, lateral (G) and frontal (H). Further abbreviations: nt: neural tube, pc: precartilage condensation.

At day 12.5 of gestation, the same activity pattern of the fusion protein was observed as in the earlier stages, with additional expression in mesencephalon, telencephalon, spinal ganglia, follicle plugs and mammary buds (Fig. 10A-F). It was not clear if the appearance of stronger staining in some tissues, e.g. areas of precartilage

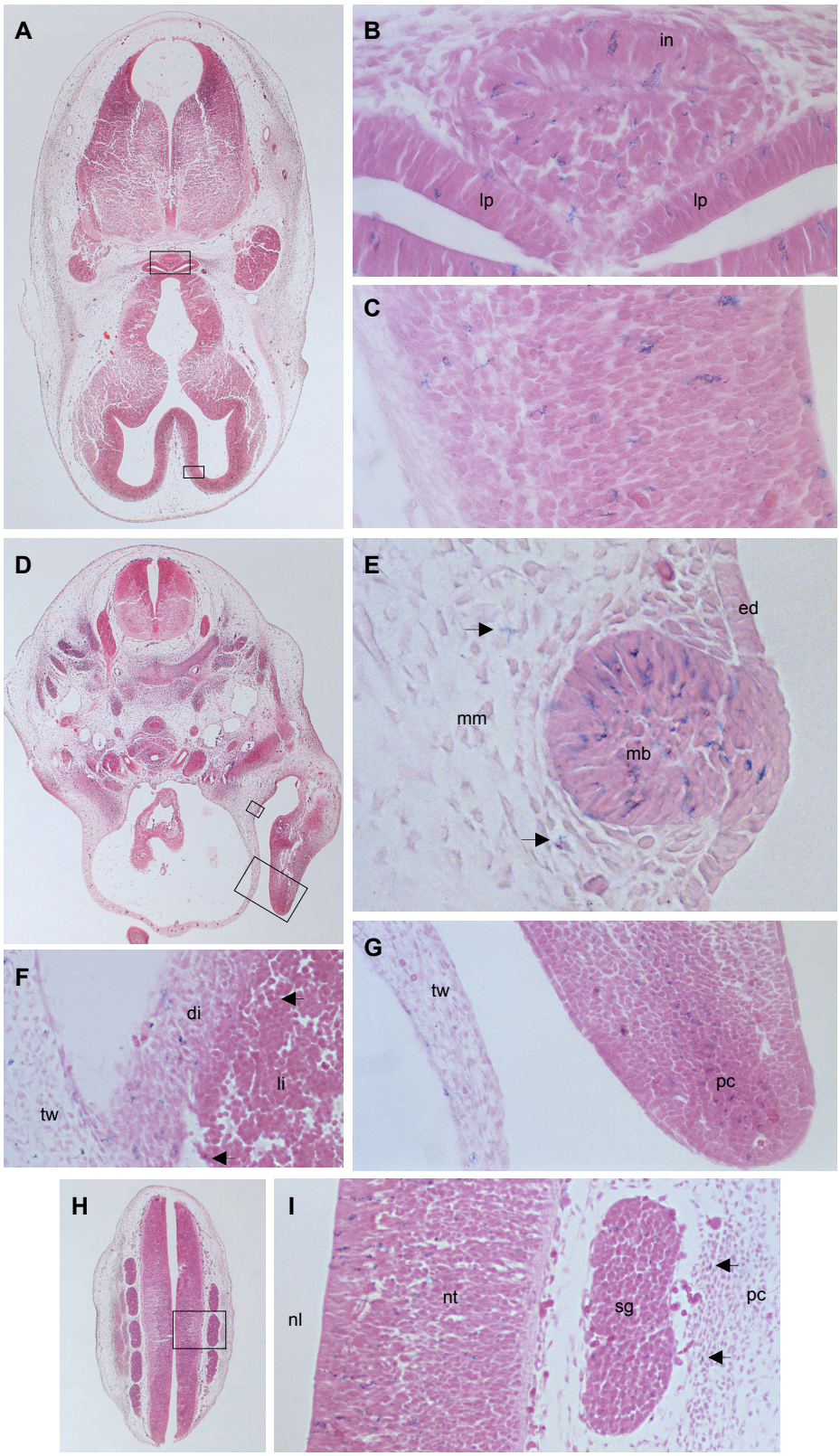


Figure 11. KRAB/β galactosidase fusion protein expression in sections of an E12.5 homozygous mutant embryo. (A) Section through head, boxes refer to magnified (B) infundibulum (in) and anterior lobes of the pituitary (lp), and (C) the neopallial cortex. (D) Section through thorax, boxes refer to magnified (E) mammary bud (mb) and (G) left forelimb with thoracic body wall (tw). Arrows in G indicate stained cells in mammary mesenchyme (mm). (F) Magnified view of the liver (li) with diaphragm (di). Arrows in F indicate stained cells in the liver. (H) Section through rump, box indicates magnified (I) neural tube (nt) and spinal ganglion (sg). Arrows in I indicate stained cells in the precartilage condensation (pc). Embryo was stained with Bluo-gal, transverse sectioned at 10 µm and counterstained with eosin. Further abbreviations: ed: epidermis, nl: neural lumen.

condensation in limb buds (Fig. 10G-H), was a result of increased expression of the fusion protein or simply regions of higher cell density. To obtain better resolution of the expression of the fusion protein, paraffin sections were made of stained E12.5 embryos. From examination of transverse sections (Fig. 11), it was apparent that the number of stained cells was similar between areas of different cell densities. Thus, areas and tissues having darker staining in whole-mounts appear due to the compactness of the tissue.

In later stages, i.e. E13.5 and E17.5, expression of the fusion protein was seen as in earlier embryos, with new regions of expression in the olfactory bulb and lung (data not shown).

The results demonstrate that the KRAB-ZF gene is expressed at least from embryonic stages E9.5 to E17.5. The expression of the fusion protein does not appear restricted to certain tissues or cell types. Staining with Bluo-gal was also conducted on isolated organs of adult mice. Unfortunately even after repeated experiments, no signal could be detected (data not shown). In this case, RT-PCR offered a more sensitive method for verifying expression of the KRAB zinc-finger gene.

3.2.3. The endogenous KRAB-ZF transcript exists as two splice variants in embryos and organs

To characterize the expression of the wild-type KRAB zinc-finger transcript, as well as to look for splice-variants, RT-PCR was conducted on embryos, adult organs, and skin from postnatal mice in the C57BL/6J background. Embryos at stages E11.5, E12.5 and E13.5 were chosen as a reference for the initiation of pigmentation in the retinal pigment epithelium of the eye. Postnatal skin was included to look for differences in gene expression around the time of beginning pigmentation in the skin.

PCR primers flanking the ORF were expected to generate a product of 2377 bp (Fig. 12A). Interestingly, RT-PCR analysis produced not only the expected fragment of 2377 bp, but also a shorter fragment of 1566 bp. Both products were expressed in all stages of wild-type embryos and postnatal skin, as well as in all adult organs analyzed (Fig. 12B-D). No PCR products were obtained in homozygous mutants.

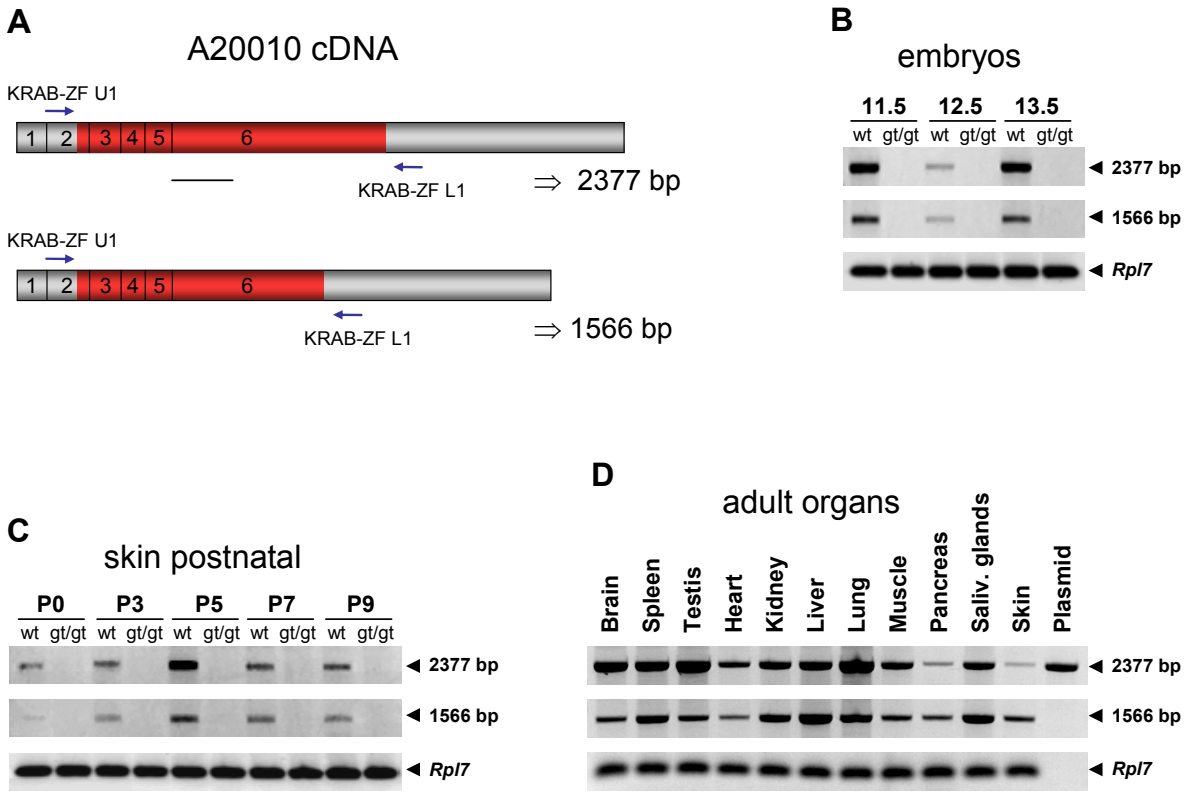


Figure 12. RT-PCR expression analysis of the KRAB-ZF transcript in mouse embryos and tissues. (A) Gene specific PCR primers (KRAB-ZF U1 and KRAB-ZF L1) were selected that flank the ORF (red) of the wild-type A20010 cDNA. A PCR product of 2377 bp is expected. The line below exon 6 represents the location of the 811 bp sequence missing in the splice variant producing the 1566 bp fragment. In wild-types, a fragment of the expected size, as well as a fragment of smaller size was detected in all embryo stages (B), skin of postnatal mice (C), and adult organs (D) tested. Neither of the fragments was detected in homozygous mutants (gt/gt). *Rpl7* was used as a quantity control, and a plasmid containing the full-length wild-type cDNA was used as a positive control.

Sequencing both fragments confirmed that the 2377 bp product contained the putative full-length ORF, whereas the second product, 1566 bp, represented a splice-variant from which the first 811 nucleotides of exon 6 were missing (Fig. 12A). This appears to be due to a cryptic splice site in exon 6 and leads to the loss of the sequence encoding the first 2 zinc-finger motifs. The new putative ORF was extended 61 nt to a new stop codon. Translation of this ORF led to a product in which the remaining 8 zinc-finger motifs and nuclear localization signal were no longer encoded, due to a frame-shift. BLAST results from the putative full-length protein indicated 100% identity to a protein encoded by the cDNA sequence BC063263, whereas results from the putative isoform indicated 100% identity to the protein encoded by BC063263 up to the new splice site. No significant identities were found with the protein sequence C-terminal to the splice site. It is therefore unclear if a protein is produced from the new ORF as predicted.

3.2.4. The A20010 mutant mouse line displayed hypopigmentation

The preliminary genetic analysis of the A20010 mouse line suggested a strong correlation between homozygous mutants and hypopigmentation of the coat and eyes (Prochnow, 2002). A few initial litters from heterozygous matings in the 129/Sv genetic background produced offspring that were brown, gray/brown, beige or cream in coat color (Fig. 13A). The parents were brown (agouti) in coat color. Brown offspring were almost exclusively wild-type or heterozygous for the KRAB-ZF mutation. Mice with a gray/brown coat color were heterozygous. Lighter coat colors, i.e. beige and white, corresponded to homozygous mutant mice. This correlation was also observed in the C57BL/6J genetic background in which black (nonagouti) heterozygous parents produced offspring with black, light gray or white coat color (Fig. 13B). Furthermore, hypopigmentation of the coat was associated with hypopigmentation of the eyes. These observations led to the initial conclusion that the A20010 KRAB-ZF protein might be involved in the development of pigment producing cells or the regulation of melanin production. To verify the preliminary observations, analysis was conducted on a larger number of offspring. In addition, eyes and hair from wild-type and homozygous mutant mice were examined for differences in melanin color and amount, as well as type of pigment producing cells.

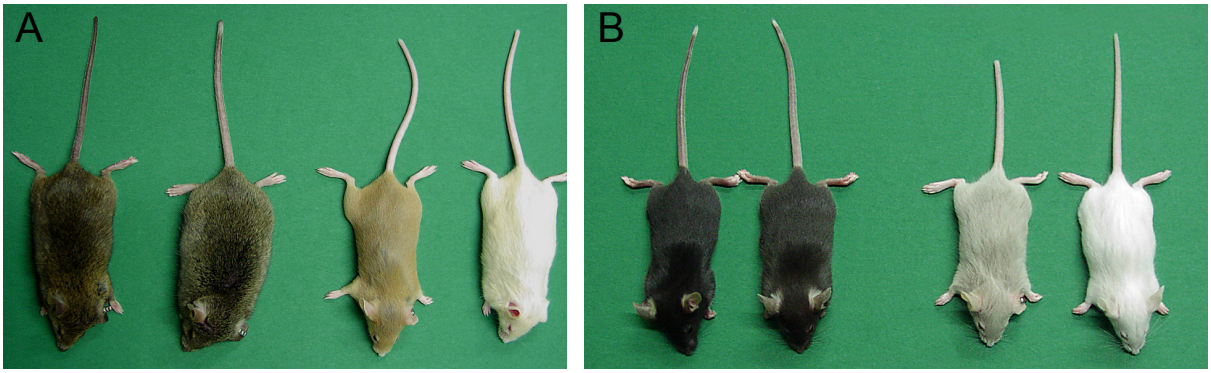


Figure 13. Photographs of mice from the A20010 mutant mouse line. (A) Offspring from heterozygous matings in the 129/Sv genetic background. Coat colors were classified from left to right as brown, gray/brown, beige, and cream. Brown mice were wild-type, heterozygous, or in seldom cases homozygous for the KRAB-ZF mutation. All gray/brown mice were heterozygous. Beige and cream colored mice were homozygous mutants. The heterozygous parents were brown (agouti) in coat color. (B) The coat colors in the C57BL/6J background were black, light gray, and white. From this single litter, one black offspring was wild-type and one heterozygous. The light gray and white mice were homozygous mutants. The heterozygous parents were black (nonagouti) in coat color.

To reconfirm the previous analysis of coat color in the mutant mouse line, a larger number of heterozygous KRAB-ZF mutant mice with wild-type coat color were mated.

Due to pigmentation differences in the 129/Sv (agouti) and C57BL/6J (non-agouti) genetic backgrounds, both mouse strains were used for independent analysis.

In the 129/Sv genetic background, 213 offspring were genotyped and analyzed for coat color (Fig. 14A). All 52 wild-type animals displayed a brown coat. From 112 heterozygous mice, 100 showed brown coat color, 9 gray/brown, 1 beige and 2 cream. From 49 homozygous mutant mice most displayed hypopigmentation. 20 homozygous mice had a beige, 24 a cream and 5 a brown coat.

A similar correlation was observed in the C57BL/6J genetic background (Fig. 14B). From 242 offspring, 55 were wild-type, all of which had a black coat. 120 mice were heterozygous with 116 black, 2 light gray and 2 white. Of the 67 homozygous mutants, 19 were light gray, 44 white and 4 black.

Mice with dark, normal pigmented coat color always had dark, normal pigmented eyes, whereas hypopigmented mice always possessed hypopigmented eyes. This indicated a strict correlation of pigmentation in both organs. These results verified cosegregation of the mutation and hypopigmented coat and eyes.

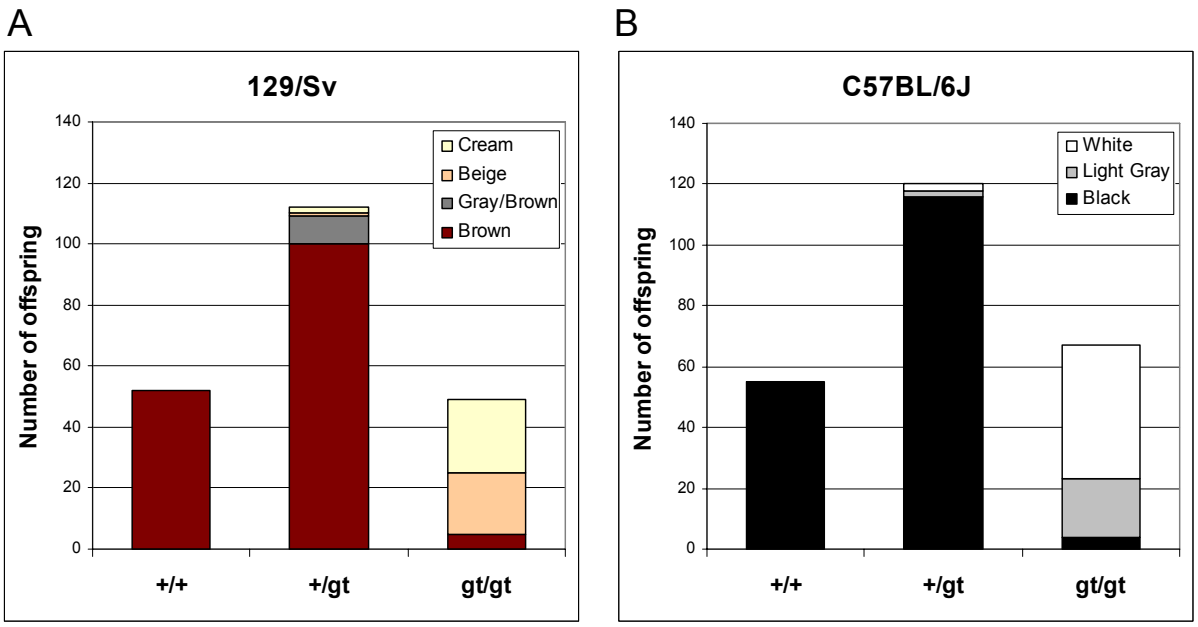


Figure 14. Coat color distribution among offspring of heterozygous matings.
(A) Distribution of coat color and genotype of 213 mice analyzed in the 129/Sv genetic background. Parental mice were brown, agouti, in coat color. (B) Distribution of coat color and genotype of 242 mice analyzed in the C57BL/6J genetic background. Parental mice were black in coat color. See also figure 13 for coat colors.

The genetic analysis also indicated that the distribution of wild-type to heterozygous and homozygous offspring in the 129/Sv background was 52 : 112 : 49 and in the C57BL/6J background 55 : 120 : 67. These values represented the normal 1 : 2 : 1 Mendelian distribution, indicating that the mutation had no effect on viability.

3.2.4.1. Hypopigmentation was found in hair follicles and shafts

Pigmentation in mice is found in the eyes, hair and to a lesser degree in the skin. In these tissues the observed color is based on the type of melanin produced and on the distribution of the melanin containing granules. There are two forms of melanin, black to dark-brown eumelanin and red to yellow pheomelanin.

To define the color and distribution of melanin in hair shafts and follicles, whole-mount skin samples were taken from C57BL/6J wild-type mice with a black coat and homozygous mutant mice with a light gray or white coat. The hair shafts and follicles of wild-type mice contained black pigmentation granules, presumably eumelanin (Fig. 15 A and D). The hair shafts and follicles of light gray homozygous mutants appeared to have similar amounts of pigment as compared to the wild-type mice, but the pigmentation was reddish-brown (Fig. 15 B and E). This suggested a partial switch from eumelanin to pheomelanin. White homozygous mutants also exhibited pigmentation, but the number of pigmented hair shafts and follicles was greatly reduced, along with the amount of pigment (Fig. 15 C and F). The color of the pigment was lighter than that of the light gray mouse. This suggested a complete switch to pheomelanin.

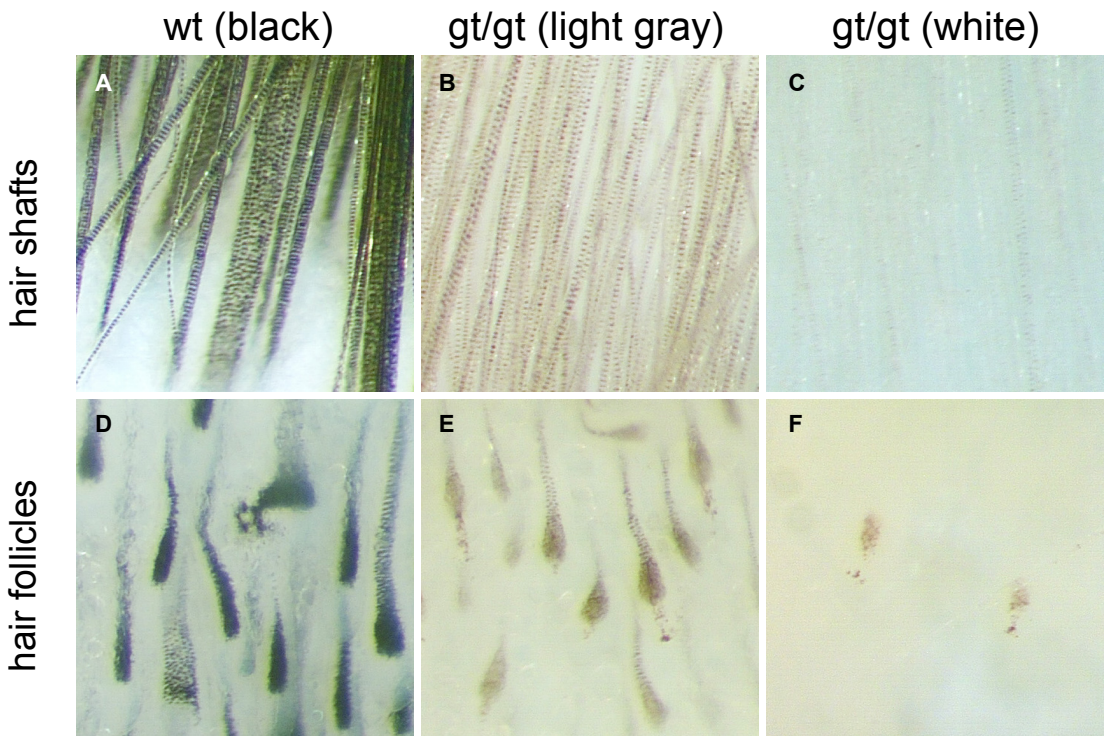


Figure 15. Pigmentation of hair shafts and follicles in wild-type and homozygous mutant mice. Skin with hair shafts and follicles from a wild-type (wt) mouse (A, D), a light gray mutant (gt/gt) mouse (B, E), and a white mutant mouse (C, F) in the C57BL/6J background. Pigmentation in the hair shafts and follicles from mutant mice suggests a switch in melanin pigment. Also, the number of pigmented hairs and follicles in the white mutant is greatly reduced, as compared to the number in wild-types.

Viewing single hairs under higher magnification, it was again observed that the wild-type hair appeared to contain exclusively black eumelanin (Fig. 16A). The homozygous mutant with light gray coat color displayed varying degrees of lighter pigmentation (Fig. 16B). This is most likely a combination of eumelanin and pheomelanin, whereas in the mutant with white coat color, only a small number of melanin granules was found (Fig. 16C arrow), but not in every hair. These were presumably pheomelanin granules.

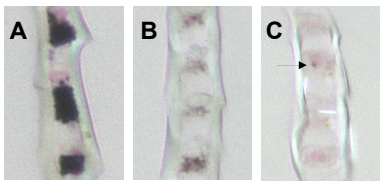


Figure 16. Magnified view of sections of single hairs from wild-type and homozygous mutant mice. (A) Hairs of C57BL/6J wild-type mice contain eumelanin in the form of black pigment granules. (B) In hairs from mutants with light gray coat color, the pigment granules are reduced in number and of a lighter color, suggesting a switch from eumelanin to pheomelanin. (C) In mutants with white coat color, the hairs are virtually devoid of pigmentation. The arrow indicates a larger melanin granule, with several smaller alongside it. Pink coloration is due to eosin counter-staining.

The observation of pigmentation in the hair follicles and shafts of mutants verified the presence of melanocytes and the lighter color of melanin. The reduction in amount and color of melanin explained the hypopigmentation. To resolve the possible cause, it was necessary to determine whether the mutant mice had a reduction in the number of melanin producing cells or a defect in melanin production. To answer this, pigmentation in the eye was characterized. Although the eye contains two different cell types that produce pigment, they use the same process to produce melanin.

3.2.4.2. Hypopigmentation was found in adult eyes of mutants

In the adult eye, pigmentation is found in the retinal pigment epithelium (RPE), choroid and iris. The iris consists of two pigmented layers: an anterior layer of iris pigmented epithelium (IPE) and a posterior layer of melanocytes. The melanocytes of the iris and choroid are of neural crest origin, whereas the RPE is derived from the outer wall of the optic cup. The IPE is an extension of the RPE. If the hypopigmentation in the A20010 mutant mouse line resulted from a reduced or missing cell type, then it was unlikely that melanocytes and pigmented epithelium would both be affected, the cells being of different origins. If the defect were subcellular (e.g. in melanosomogenesis or melanogenesis), the epithelial cells, as well as the melanocytes would display a reduction or absence of pigmentation. To determine the defect, whole-mount eyes and sections were analyzed.

The eyes from wild-type mice displayed black pigmentation of the iris and RPE (Fig. 17A). Pigmentation was also found in the eyes of mutant light gray mice, where the amount and color were obviously reduced (Fig. 17B). As was observed in the hair shafts and follicles, the melanin had a reddish-brown color. No pigment was found in the eyes of white mice (Fig. 17C).

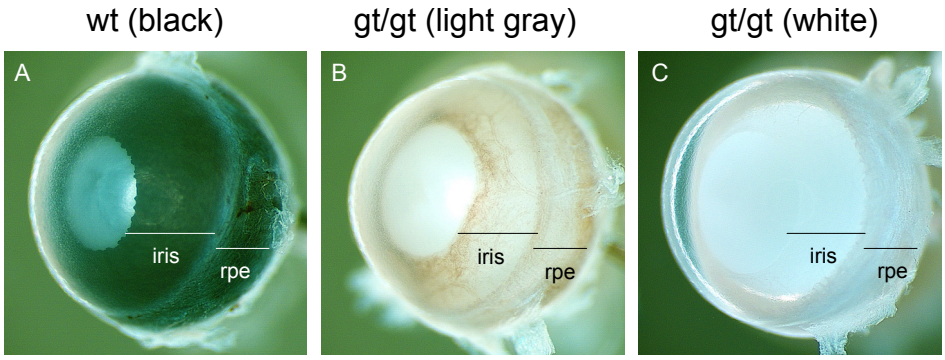


Figure 17. Pigmentation in whole-mount eyes of adult wild-type and homozygous mutant mice. (A) Normal pigmentation seen in the eye of a wild-type mouse. (B) Reduced pigmentation found in the iris and RPE of a homozygous mutant mouse with light gray coat color. (C) No pigmentation displayed in the eye of a homozygous mutant mouse with white coat. The pink color of the retina is due to vascularization. Lines indicate the extent of the iris and retinal pigment epithelium (rpe).

Because the iris consists of a layer of pigmented epithelium and a layer of melanocytes, and the RPE is covered by the melanocyte-infiltrated choroid, it was not possible to determine the precise source of pigmentation in the whole-mount eyes. Therefore, paraffin sections were made and stained with hematoxylin and eosin. In the eyes of wild-type mice, pigmentation was found in both layers of the iris, as well as in the ciliary body and RPE (Fig. 18 A, D and G). Pigmentation was reduced in these structures in light gray homozygous mutants (Fig. 18 B, E and H) and completely missing in the white mutant (Fig. 18 C, F and I). Very weak pigmentation was found in the choroid of wild-type and light gray mutant mice.

The reduction or absence of pigmentation in different melanin producing cells of distinct lineages, i.e. melanocytes and pigment epithelium, indicated that the hypopigmentation phenotype was not a result of a cell type-specific defect, but rather a subcellular or biochemical defect in melanin production. To look for the gene or genes leading to such a defect, RT-PCR was conducted to characterize the expression of important pigmentation genes in wild-type and mutant mice.

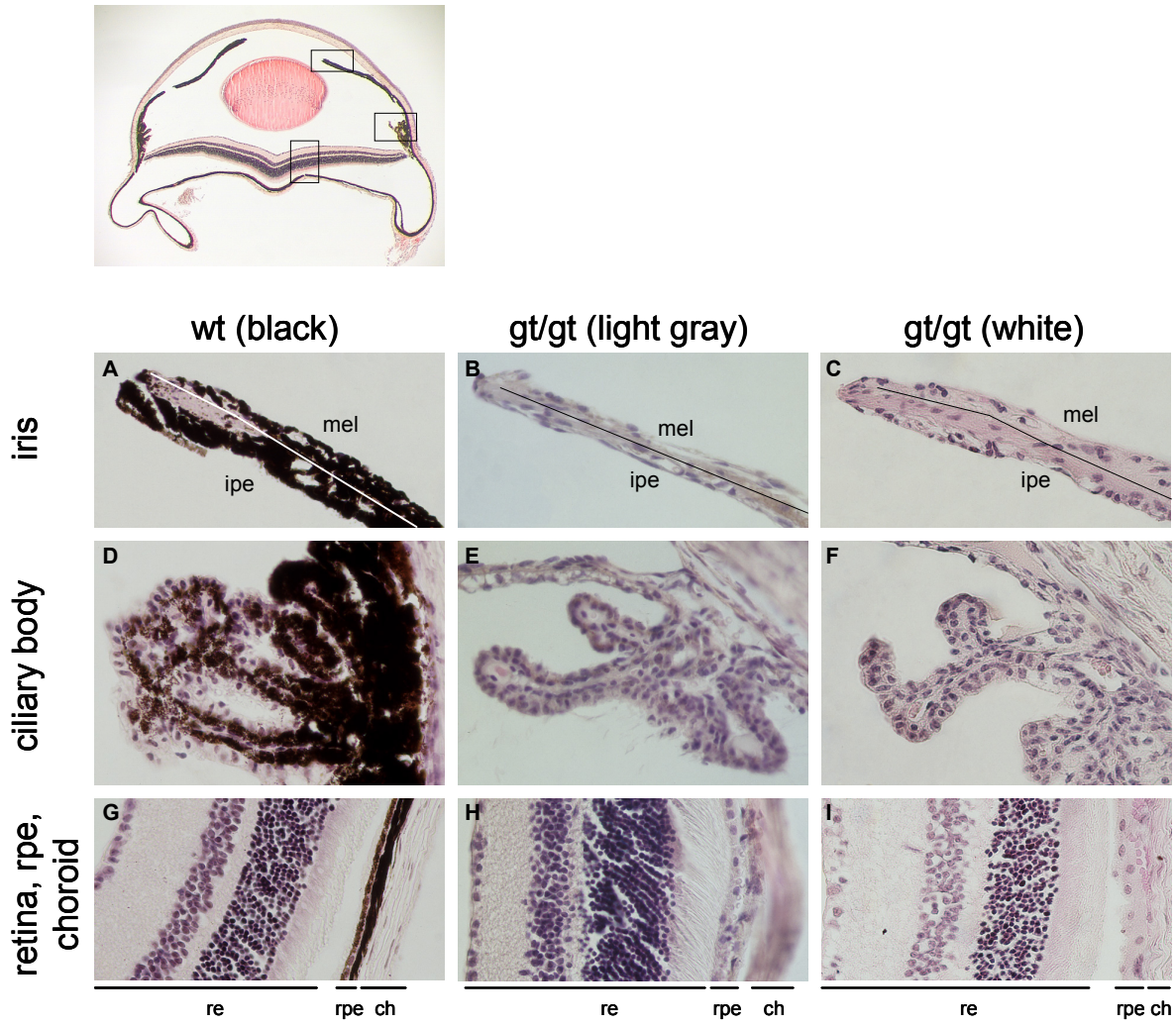


Figure 18. Sections through the eye of adult wild-type and homozygous mutant mice. Normal pigmentation is seen in the iris pigment epithelium (ipe) and melanocyte layer (mel) of the iris (A), in the ciliary body (D), and in the retinal pigment epithelium (rpe) (G) of adult wild-type (wt) mice in the C57BL/6J background. Greatly reduced pigmentation is found in the iris (B), ciliary body (E) and rpe (H) of homozygous mutants (gt/gt) having a light gray coat color. Pigmentation appears absent in the iris (C), ciliary body (F) and rpe (I) of homozygous mutants having a white coat. A line is drawn through the iris to divide the anterior iris pigment epithelium from the posterior melanocyte layer. 10 µm sections stained with hematoxylin and eosin. Further abbreviations: ch: choroid, re: retina.

3.2.4.4. *Pink-eyed dilution* and *tyrosinase* were differentially expressed in the A20010 mutant mouse line

Since the trapped KRAB-ZF gene encoded a putative transcription factor, to determine if the protein regulated the expression of genes involved pigmentation, the effect of its absence on transcript levels of pigmentation genes was analyzed by RT-PCR. This was conducted using primers for key factors involved in melanin synthesis and eumelanin-pheomelanin switching. To reconfirm the presence of pigment

producing cells, several genes were included that influence the development and migration of melanocyte precursors, i.e. melanoblasts (Table 4). The same cDNA pools were used as for the analysis of the endogenous KRAB-ZF transcript.

Development of melanoblasts in neural crest
- *paired box gene 3 (Pax3), splotch*
- *endothelin receptor type B (Ednrb), piebald spotting*

Migration of melanoblasts
- *kit oncogene (Kit), dominant white spotting*

Differentiation into melanocytes
- *microphthalmia-associated transcription factor (Mitf)*

Melanosome function
- *pink-eyed dilution (p)*

Melanin biosynthesis
- *solute carrier family 45, member 2 (Slc45a2), underwhite*
- *tyrosinase (Tyr), albino*
- *tyrosinase-related protein 1 (Tyrp1), brown*
- *dopachrome tautomerase (Dct), slaty*
- *silver (Si)*

Switch between eumelanin and pheomelanin
- *agouti (A)*
- *pro-opiomelanocortin-α (Pomc1)*
- *melanocortin 1 receptor (Mc1r), extension*
- *attractin (Atrn), mahogany*
- *mahogunin, ring finger 1 (Mgrn1)*
- *gamma-glutamyltransferase 1 (Ggt1)*
- *osteopetrosis assoc. transmembrane protein 1 (Ostm1), grey-lethal*

Table 4. Marker genes in the pigmentation pathway. The genes are listed by their current name, with symbol in parentheses, followed by the allele associated with a pigmentation phenotype, if the allele name is different than gene name. Primers for these genes were used for RT-PCR analysis.

The majority of genes showed relatively equal transcript levels in wild-type and mutant embryos (Fig. 19). This included *Microphthalmia-associated transcription factor* (*Mitf*), which was used as a marker for differentiating melanocytes (Fig. 20A). This suggested that there was no reduction in melanocyte numbers, which added to the evidence that the hypopigmentation was not due to loss of pigment producing cells. Only transcript levels of *pink-eyed dilution* (*p*) and *tyrosinase* (*Tyr*) were reduced in mutant embryos.

Gene	wt	gt/gt
Pax3	++	++
Ednrb	++	++
Kit	++	++
*Mitf**	++	++
p*	**++**	**+**
Slc45a2	++	++
Tyr*	**++**	**+**
Tyrp1	++	++
Dct	++	++
Si	++	++
A	-	-
Pomc1	++	++
Mc1r	++	++
Atrn	++	++
Mgrn1	++	++
Ggt1	++	++
Ostm1	++	++

+ lower expression
++ higher expression
- not expressed

Figure 19. Expression of pigmentation marker genes in wild-type and homozygous mutant embryos. Results of RT-PCR analysis for pigmentation genes with cDNA from wild-type (wt) and homozygous mutant (gt/gt) mice at embryonic stages E11.5, E12.5 and E13.5. The *pink-eyed dilution* (*p*) and *tyrosinase* (*Tyr*) genes (in bold) exhibited the most significant variation between wild-type and mutant. The (*) indicates genes with results shown in figure 20. No expression of the *agouti* transcript was detected, as the mice were C57BL/6J.

Pink-eyed dilution, which is necessary for proper functioning of the melanosome, showed a clear reduction of transcript levels at stage E11.5 in homozygous mutants (Fig. 20A). At stages E12.5 and E13.5 levels did not appear significantly different as in wild-type. Since pigmentation of the eye is first obvious around E11.5 it might be that the reduction of *p* at this stage is sufficient to prevent proper melanin production.

Tyrosinase, which is the rate-limiting enzyme in melanin biosynthesis, exhibited the most interesting results. For the detection of *tyrosinase* expression the expected product size was 438 bp (Fig. 20B). This product was found strongly reduced in mutant embryos. In addition, a second, smaller PCR product of 290 bp was detected which displayed increased expression in mutants (Fig. 20C).

To characterize *tyrosinase* expression at the initiation of pigmentation in the hair follicles and shafts, RT-PCR was also conducted on skin taken from several postnatal stages ranging from P0 to adult. As in the embryos, both PCR products were found in the skin. Again, in homozygous mutants the expression of the expected product was reduced, whereas expression of the smaller product was increased as compared to wild-type (Fig. 20D).

Both of the *tyrosinase* PCR products were isolated and sequenced. The larger fragment of 438 bp corresponded to the expected sequence using the full-length cDNA as a reference. The smaller fragment of 290 bp represented a splice-variant in which exon 3 had been removed. This splice-variant is known as Δ3 and has been reported as producing an inactive enzyme (Porter and Mintz, 1991; Terao *et al*., 1989). In mutants, a switch to this inactive form seems to occur.

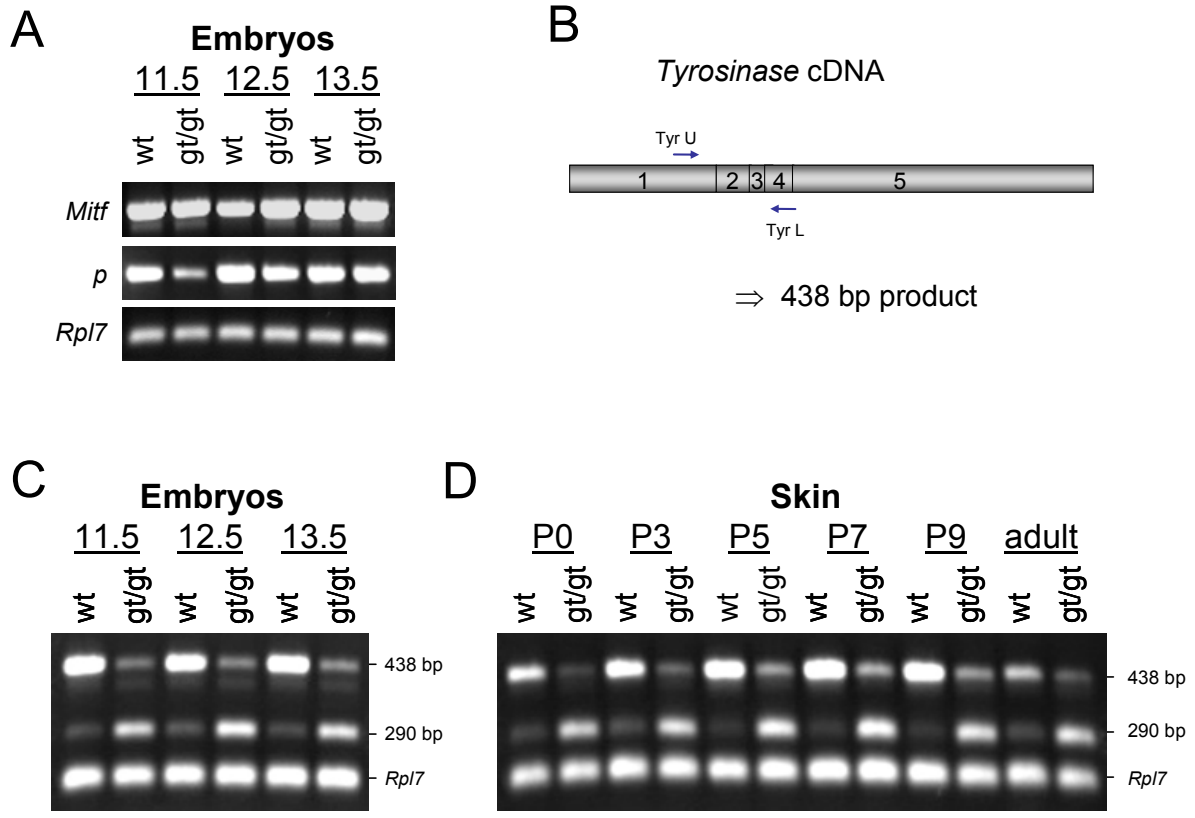


Figure 20. Gel electrophoresis of products from RT-PCR using pigmentation gene specific primers. cDNAs of RNAs from wild-type (wt) and homozygous mutant (gt/gt) embryos and postnatal skin were analyzed by RT-PCR. (A) *pink-eyed dilution* (*p*) and (C) *tyrosinase* (*Tyr*) were the only genes to show a reduction in transcript levels in mutant embryos. (B) Location of the primers used with respect to the *Tyr* cDNA, and the expected fragment size of 438 bp. (C-D) RT-PCR with *tyrosinase* primers produced not only a product of the expected size of 438 bp, but also a smaller product of 290 bp. Expression of *Mitf* in A indicates that the melanocyte numbers in mutant embryos is not reduced. *Rpl7* is used as a quantity control.

Results from the RT-PCR analysis clearly indicated that the hypopigmentation phenotype seen in the A20010 gene trap mouse line is a result of a combined reduction in functional tyrosinase and pink-eyed dilution. Mutations of these two genes have already been known to cause hypopigmentation in mice and humans.

3.2.5. Mutant alleles of *pink-eyed dilution* and *tyrosinase* cosegregated with the KRAB-ZF gene trap allele

During the finalization of the RT-PCR analysis, improvements to the sequence alignment of the Mouse Genome had allowed for chromosomal assignment of the putative KRAB-ZF gene. This was found to be on chromosome 7 at the genetic map position of 20 cM. Using the International Federation of Pigment Cell Societies database to search for genes involved in pigmentation, it was found that the *pink-eyed dilution* and *tyrosinase* loci were also located on chromosome 7. The *p* locus is located at the 28 cM position and the *Tyr* locus at the 44 cM position (Fig. 21).

For the production of the A20010 gene trap line, the hybrid ES-cell line R1 had been used. From literature, it was found that R1 ES-cells are heterozygous for the *p* and Tyr^{ch} (chinchilla) mutant alleles, with one chromosome carrying both mutations and the other chromosome being wild-type (Nagy *et al*., 1993; Simpson *et al*., 1997; Threadgill *et al*., 1997). To determine if these mutations were cosegregating on the same chromosome with the KRAB-ZF gene trap mutation, the genotype at the *p* and *Tyr* loci were experimentally verified.

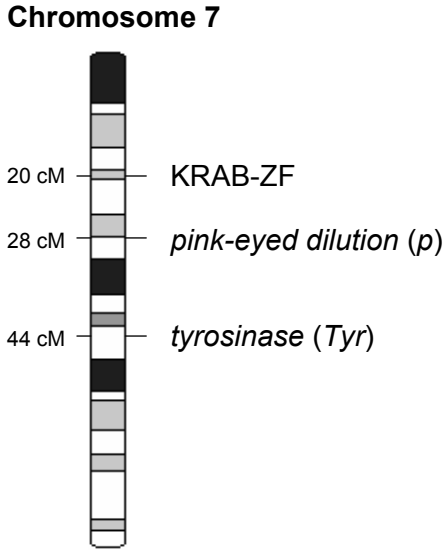


Figure 21. Schematic diagram of mouse chromosome 7.
The putative KRAB-ZF locus is located at the genetic map position of 20 cM, whereas the *p* and *Tyr* loci are located at the positions 28 cM and 44 cM, respectively.

The strategy for genotyping the *tyrosinase* locus was based on the findings of Beermann *et al*., (1990) who reported the chinchilla mutant allele as being a single nucleotide substitution in the wild-type allele. The substitution of a guanine for adenine at position 1505 of the *tyrosinase* mRNA leads to an amino acid exchange of alanine for threonine. This causes an increased susceptibility of the protein for proteolytic degradation (Halaban *et al*., 1988). After sequencing the site of substitution, wild-type mice were found to carry the wild-type allele, whereas hypopigmented homozygous mutant mice carried the substitution as mentioned above, i.e. the chinchilla allele (Fig. 22). This verified the A20010 line as not only carrying the gene trap mutation, but also the *tyrosinase* chinchilla mutation.

wild-type:	GGG		GCA		**G**CA		CTG		GTG
	Gly	—	Ala	—	**Ala**	—	Leu	—	Val
chinchilla:	GGG		GCA		**A**CA		CTG		GTG
	Gly	—	Ala	—	**Thr**	—	Leu	—	Val

Figure 22. The chinchilla mutation at the *tyrosinase* locus.
The chinchilla allele differs from the wild-type allele by a single nucleotide substitution of a guanine (G) for an adenine (A) in exon 5. This leads to an amino acid exchange of alanine (Ala) for threonine (Thr). In homozygous form, this mutation leads to hypopigmentation through a strong reduction in tyrosinase. Sequencing was conducted with gene specific primers Tyr U2 and Tyr L2

To determine if the A20010 line was also carrying a mutation at the *p* locus, genotyping was conducted as first described by Brilliant *et al*., (1994). This consisted of PCR analysis of genomic DNA using primers that detect a polymorphism in intron 6 of the *p* locus. Wild-type mice produce a 423 bp fragment and mutant mice a 470 bp fragment. This analysis was conducted on mice of both genetic backgrounds, 129/Sv and C57BL/6J, with all three genotypes at the KRAB-ZF locus, and with either a normal or hypopigmented coat color (Fig. 23). In the 129/Sv genetic background the wild-type mouse with brown coat color displayed only the 432 bp fragment, as expected. The gene trap mutant mice, on the other hand, all exhibited a 470 bp mutant *p* fragment, thus indicating the presence of the *p* mutation in the A20010 line. The gene trap heterozygote, with wild-type coat color, was heterozygous for the *p* mutation. The homozygous mutant with beige coat color, which is an intermediate form of hypopigmentation, showed a yet uncharacterized fragment of intermediate size. The homozygous mutant with the stronger

hypopigmentation displayed only the mutant *p* fragment. Interestingly, the homozygous mutant with wild-type coat color exhibited mutant and wild-type *p* fragments. In this case, the recessive nature of the *p* mutation is responsible for the wild-type coat color.

Virtually identical results were found in the C57BL/6J background (Fig. 23). The wild-type gene trap mouse was wild-type at the *p* locus, the heterozygous gene trap mouse was heterozygous at the *p* locus, the homozygous mutant with intermediate hypopigmentation displayed mutant and intermediate *p* fragments, and the homozygous mutant with wild-type coat color was heterozygous at the *p* locus. The only difference to the 129/Sv background was that the homozygous mutant with the strongest hypopigmentation (white) also exhibited an intermediate *p* fragment.

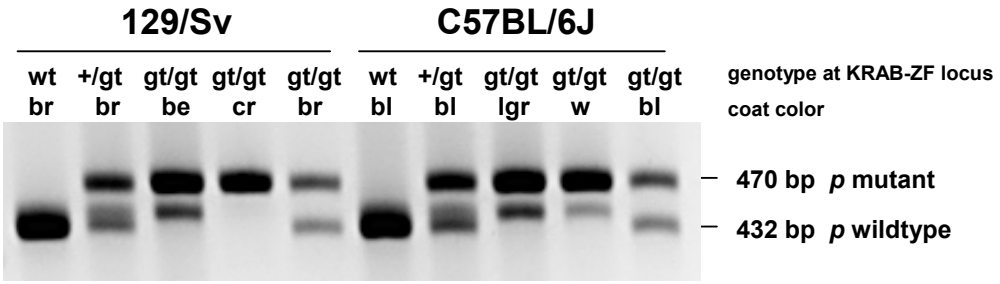


Figure 23. Genotyping mice of the A20010 gene trap line at the *p* locus.
(A) Mice carrying the gene trap mutation in the 129/Sv and C57BL/6J genetic backgrounds were also found to be carrying a mutation at the *p* locus. Coat colors: be: beige, bl: black, br: brown, cr: cream, lgr: light gray, w: white. Genotyping was conducted with gene specific primers p U2 and p L2.

Genotyping the *tyrosinase* and *p* loci showed that the A20010 gene trap mouse line was carrying mutations at both loci, which cosegregated with the trapped gene. It was therefore clear that the *Tyr* and *p* mutations were responsible for the hypopigmentation phenotype. This raised the question of the cause of wild-type pigmentation (i.e. black in C57BL/6J and brown in 129/Sv) and intermediate hypopigmentation (i.e. light gray in C57BL/6J and beige in 129/Sv) of some of the homozygous mutants. Mutations of *pink-eyed dilution* are associated with reductions in pigmentation and a partial shift from eumelanin to pheomelanin, whereas mutations of *tyrosinase* are associated with complete or almost complete absence of pigmentation (Silvers, 1979). It was therefore likely that wild-type pigmentation in homozygous mutants was caused by recombination of mutant *p* and *Tyr* alleles for

wild-type alleles by chromosomal crossing-over somewhere between the KRAB-ZF locus and the *p* locus. The distance between these two loci is 8 cM (Fig. 21), indicating approximately 8% occurrence of recombination. In C57BL/6J, 4 of the 67 homozygous mutants had wild-type pigmentation. This corresponds to a 6% frequency of occurrence or a 6% recombination for the area between the KRAB-ZF and *p* loci. In 129/Sv, 5 of the 49 homozygous mutants had wild-type pigmentation. This corresponds to a 10% frequency. In the case of homozygous mutants with intermediate hypopigmentation, it is likely that recombination occurs somewhere between the *p* locus and the *Tyr* locus. Presence of only the *p* mutation would lead to the light gray and beige coat colors. The distance between the *p* and *Tyr* loci is 16 cM (Fig. 21), which represents a 16% frequency of recombination. In C57BL/6J, 19 of the 67 homozygous mutants were light gray. This corresponds to a 28% recombination frequency. In 129/Sv, 20 of the 49 homozygous mutants were beige. This represents a frequency of 41%. These results support the hypothesis that wild-type and intermediate pigmentation in homozygous mutants is a result of recombination at the *p* and *Tyr* loci.

To analyze the effect of the KRAB-ZF mutation, it was necessary to reestablish the A20010 line without the *Tyr* and *p* mutations. In the C57BL/6J background, wild-type mice were crossed with gene trap homozygous mutants displaying a wild-type coat color and heterozygosity at the *p* locus (Fig. 24, parents). If recombination had occurred at the *p* locus, it was likely that the *tyrosinase* locus would also be affected. All offspring were heterozygous for the gene trap mutation, black in coat color and either wild-type or heterozygous for the *p* mutation (Fig. 24, offspring). The offspring that were wild-type at the p locus were used for reestablishing the A20010 line, as well as intercrossed for analyzing gene trap homozygous mutants. This procedure was also conducted in the 129/Sv genetic background (data not shown).

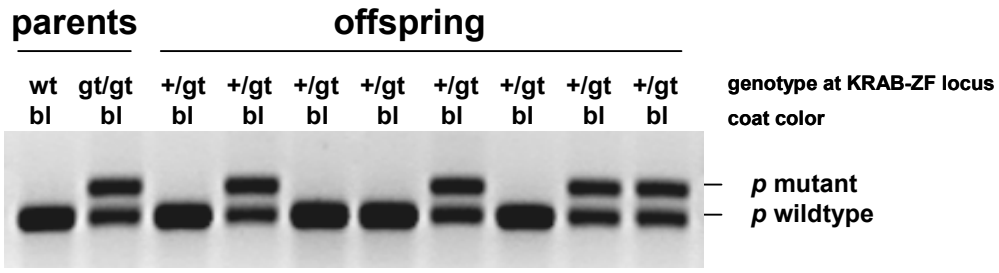


Figure 24. Genotyping mice of the A20010 gene trap line at the *p* locus.
A cross between a wild-type mouse and a homozygous mutant gene trap mouse with wild-type coat color (parents) resulted in heterozygous offspring with black coats (bl). One half were wild-type at the *p* locus, the other half heterozygous (offspring). Genotyping was conducted with gene specific primers p U2 and p L2.

Homozygous mutants in the reestablished A20010 line displayed no obvious phenotype. Analysis of weight, general appearance and behavior, as well as morphology of organs indicated no difference in mutants as compared to wild-type mice. Because phenotypes sometimes become more obvious in aged mice, the analysis should be extended to older mice.

3.3. Analysis of the gene trap mouse line A008A01

At the beginning of the work with the A008A01 mouse line, the trapped gene was completely unknown. The uncharacterized YjeF-N domain was the only identifiable domain in the protein. Analysis of the β-galactosidase activity of the fusion protein in mutant mice had shown that the gene is widely expressed throughout the embryo and the adult mouse (Meyer, 2003; Garreis, 2005). No expression was found in embryonic liver and adult heart. It was also observed that homozygous mutants were lighter and had less body fat. In some mutants, reduced “latency to fall” in the Hanging-Wire Test was observed. This might be due to reduced muscle strength. The analysis here was therefore focused on weight, fat, and muscle of the mutant mice.

3.3.1. Site of vector integration in the A008A01 trapped gene

In the initial analysis of the A008A01 mutant mouse line it was found that the PT1βgeo vector (Fig. 25A) integrated into intron 4 of the trapped gene (Fig. 25B) (Meyer, 2003). This locus is 42.3 kb in size and consists of 7 exons. The mRNA is 3327 bp long with a predicted open reading frame of 1527 bp. This ORF codes for a protein of 508 amino acids with a calculated mass of 55.9 kDa.

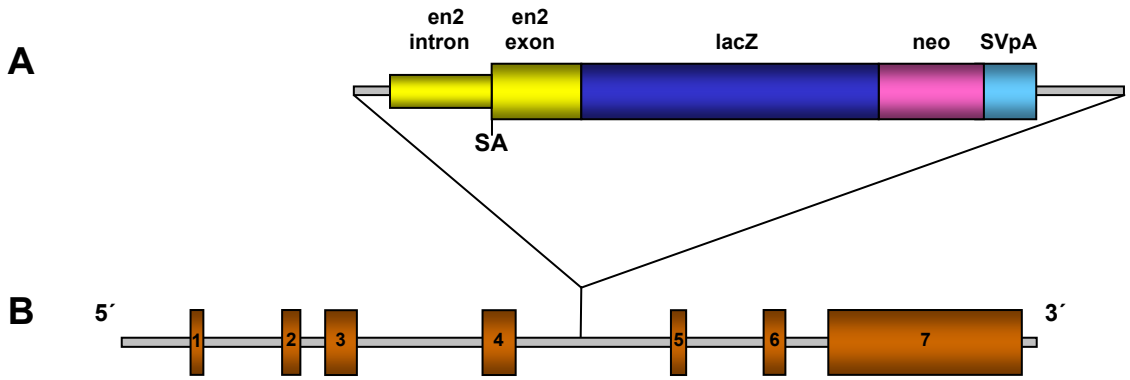


Figure 25. Schematic representation of the PT1βgeo splice-acceptor vector and the site of its integration in the novel locus. (A) The PT1βgeo vector with *βgeo* cassette. (B) The vector has integrated into intron 4 of the locus in the mouse line A008A01. This locus consists of 7 exons (brown). en2: *engrailed2*, SA: splice-acceptor site, lacZ: *β-galactosidase* gene, neo: *neomycin phosphotransferase* gene, SVpA: simian virus polyadenylation signal. (Adapted from Meyer, 2003).

3.3.2. *Lsm16* is the gene trapped in the A008A01 mouse line

Although important sources of bioinformation (e.g. NCBI and Swiss-Prot databases) are being constantly updated, only limited data was available for the trapped gene. Additional information pertaining to identity and domain structure of the protein encoded by the trapped gene became available for the very first time during the final writing of this thesis.

According to recent entries in the Jackson Laboratories Mouse Genome Informatics (MGI) database in February 2006, the trapped gene has been officially named *LSM16 homolog* (EDC3, *S. cerevisiae*). The name identifies the gene as the ortholog of the *Saccharomyces cerevisiae* gene *Edc3*. The official symbol is *Lsm16*, with the synonym *Yjdc*.

Sequence analysis using the NCBI Conserved Domain Architecture Retrieval Tool (CDART), which had previously located only the C-terminal YjeF-N domain, amino acids 309-466, now also indicates the presence of the Lsm16-N domain. This second domain spans amino acids 2 to 63 in the 508 amino acid protein (Fig. 26) and shares 75.8% identity with the consensus sequence of the eukaryotic Lsm16-N domain. The Lsm16 protein belongs to a family of like-Sm (Lsm) proteins that associate with RNA to form the core of ribonucleoprotein particles involved in a variety of RNA processing events including pre-mRNA splicing and mRNA degradation (reviewed in Khusial *et al*., 2005).

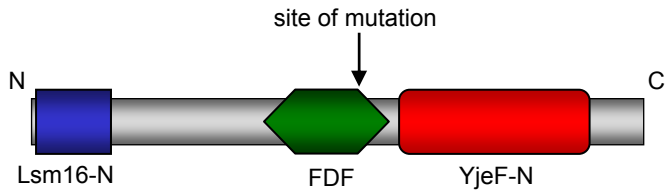


Figure 26. Schematic representation of the wild-type protein mutated in the A008A01 mouse line. Analysis of the 508 amino acid novel protein has led to the identification of 3 highly conserved domains: Lsm16-N (blue), FDF (green) and YjeF-N (red). The arrow indicates the site of mutation, which corresponds to the location of β-galactosidase and neomycin phosphotransferase sequences in the mutant fusion protein. The section of the wild-type protein C-terminal to the site of mutation is absent in the mutant fusion protein.

From comparative sequence analysis of *Lsm16* orthologs, a third domain was located, the FDF (or DFDF) domain, which spans amino acids 198 to 301 (Fig. 26). FDF, which derives its name from these three amino acids in its highly conserved consensus sequence, is yet uncharacterized. It is suggested, however, that it interacts with RNA or highly charged peptides that are commonly found in ribonucleoprotein complexes (Anantharaman and Aravind, 2004; Albrecht and Lengauer, 2004). In the A008A01 line, the gene trap vector has integrated between exons that encode the FDF domain (Fig. 26 arrow), implying that the Lsm16-N domain and three-fourths of the FDF domain are present in the Lsm16/β-gal fusion protein, whereas the YjeF-N domain is lost. The function of the YjeF-N domain is yet unknown.

Recent results from BLASTing the mouse Lsm16 amino acid sequence illustrated identities from 95.1% to 95.7% with the mammals *Bos taurus*, *Canis familiaris*, *Macaca fascicularis*, *Pongo pygmaeus*, and *Homo sapiens*. Comparing human and mouse, the Lsm16-N, FDF and YjeF-N domains shared 98.4%, 93.3% and 98.1% identity, respectively. Protein identities of the mouse sequence to *Danio rerio*, *Xenopus laevis* and *Gallus gallus* were 67.8%, 74.3% and 87.9% respectively.

Although the *Lsm16* gene has not yet been characterized in the mouse beyond this thesis and the Diploma work from Meyer (2003) and Garreis (2005), the placement of the Lsm16 protein into the family of like-Sm proteins, as mentioned above, suggests a role in RNA metabolism or processing.

3.3.3. Homozygous mutants display reduced weight

Preliminary results by others had suggested that A008A01 homozygous mutants are generally lighter than wild-type mice. To verify this observation, juvenile (6 weeks of age) and young adult mice (10 weeks of age) were examined with respect to total body weight, and weight of individual tissues.

Both in males and females, a significant reduction in weight was found between wild-type and homozygous mutant mice. At 6 weeks of age, wild-type males had an average body weight of 23.2 ± 1.1 g (mean ± s.d., n=18), whereas homozygous

mutant males had a weight of 20.0 ± 1.1 g (n=20) (Fig. 27A). This difference represents a 13.8% lower body weight in mutants. At 10 weeks of age, wild-type males had a body weight of 28.7 ± 1.4 g, whereas mutants had a weight of 25.4 ± 1.2 g. This indicates an 11.5% lower body weight in homozygous mutants. In females, body weights of 17.6 ± 1.2 g (n=21) and 15.6 ± 0.6 (n=21) for wild-types and homozygous mutants, respectively, were recorded at 6 weeks of age, and 21.0 ± 1.2 g and 19.1 ± 1.1 g for wild-types and mutants, respectively, at 10 weeks of age (Fig. 27B). This corresponds to a lower body weight in homozygous mutants of 11.4% at 6 weeks and 9.0% at 10 weeks of age.

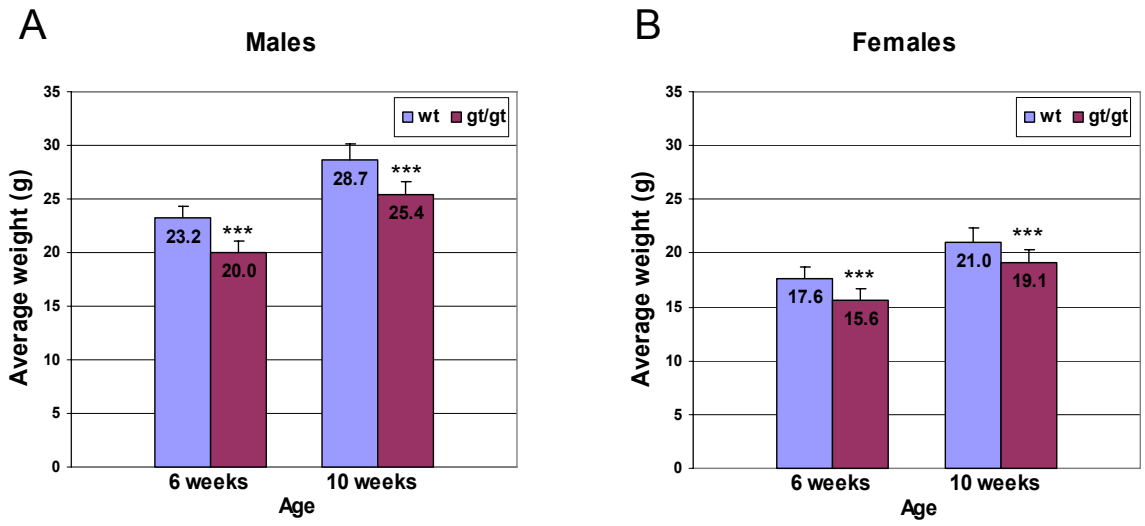


Figure 27. Body weight difference between wild-type and homozygous mutant mice. (A) Among males, a lower body weight in homozygous mutants (gt/gt) was found as compared to wild-types (wt). This was seen as 13.8% lower in juvenile mice at 6 weeks of age and 11.5% lower in young adult mice at 10 weeks of age. Males n = 18 wt, 20 gt/gt. (B) Among females, the reduction was 11.4% and 9.0% at 6 and 10 weeks, respectively. Females n = 21 for each genotype. Weights are given as mean ± s.d. ***p < 0.001.

At 20 weeks of age, the difference in body weight between wild-type (n=3) and homozygous mutant males (n=4) had increased to an average of 18.7% (Tab. 5). Moreover, the average weight of brown adipose tissue (BAT), white adipose tissue (WAT), liver and brain were 50.0%, 30.0%, 13.2% and 2.3% lower in mutants, respectively.

	wild-type (g)	gt/gt (g)	average difference (g)	(%)
Body weight	33.7 ± 1.65	27.4 ± 1.62	6.3**	18.7
BAT	0.12 ± 0.02	0.06 ± 0.01	0.06*	50.0
WAT	1.28 ± 0.18	0.90 ± 0.14	0.38*	30.0
Liver	1.82 ± 0.06	1.58 ± 0.15	0.24*	13.2
Brain	0.44 ± 0.01	0.43 ± 0.01	0.01NS	2.3

Table 5. Difference in body and tissue weights between wild-type and homozygous mutant males at 20 weeks of age. Mutant mice exhibited a lower average body and tissue weight as compared to wild-type littermates. Weights are given as mean ± s.d. n=3 wt, 4 gt/gt.
**p<0.01, *p<0.05, NS not statistically significant.

Within age-matched groups, the largest difference in overall body size between wild-type and mutant littermates was seen at 2 years of age (Fig. 28A). Although this was noticed in several wild-type and mutant mice, due to limited numbers, only one pair was further analyzed. The reduction in size of the mutant mouse was accounted for by a 22.5% decrease in head to base of tail length and a 40.1% decrease in weight (Tab. 6). After removing the skin, a marked reduction of white adipose tissue was evident in homozygous mutant compared to wild-type mice (Fig. 28B arrows).

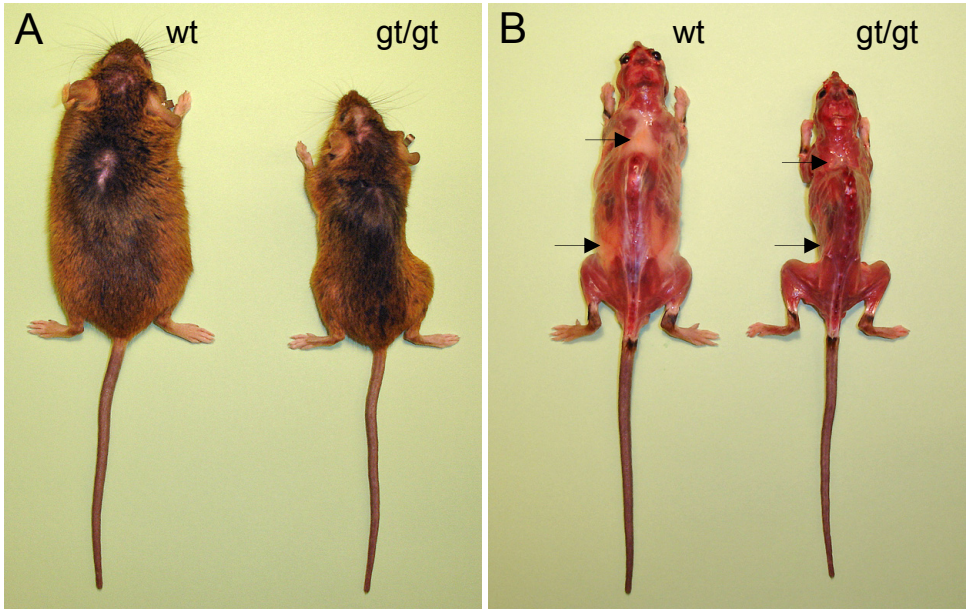


Figure 28. Difference in size and fat deposits between a wild-type and a mutant mouse at two years of age. (A) Two year old littermates in the 129/Sv genetic background. The homozygous mutant mouse (gt/gt) displays a reduction in overall body size compared to the wild-type (wt). (B) In addition, the mutant mouse is almost devoid of subcutaneous white adipose tissue (arrows).

However, the reduction in weight was found in all tissues, although to varying degrees. The largest effect was found in WAT with a 93.5% decrease, whereas the smallest reduction was seen in brain with only 4.3% (Tab. 6).

Parameter	Genotype		Reduction in gt/gt
	wt	gt/gt	
Length (cm)	8.9	6.9	22.5 %
Weight (g)	32.9	19.7	40.1 %
Tissue weight (g)			
WAT	2.30	0.15	93.5 %
Skin	5.10	1.70	66.6 %
Spleen	0.02	0.01	50.0 %
Kidney	0.49	0.25	49.0 %
BAT	0.14	0.08	42.9 %
Liver	1.26	0.82	34.9 %
Heart	0.17	0.13	23.5 %
Ga/So muscle	0.20	0.16	20.0 %
Brain	0.47	0.45	4.3 %

Table 6. Body parameters of a wild-type and a homozygous mutant mouse at two years of age. In addition to the reduction in length and weight, all tissues analyzed displayed a reduction in weight in the homozygous mutant (gt/gt) as compared to wild-type (wt). WAT: white adipose tissue, BAT: brown adipose tissue, Ga/So: gastrocnemius/soleus.

Taken together, the results demonstrate that mutant mice from 6 weeks to 2 years of age exhibited an increasing loss of body weight compared to wild-type mice. The most significant difference in individual tissues was seen in WAT, dropping from 30.0% reduction in homozygous mutants at 20 weeks to 93.5% loss at 2 years of age.

3.3.4. Endogenous *Lsm16* transcript is detected in adult tissues

Northern blot results from Meyer, 2003 showed *Lsm16* transcripts in adult brain, intestine, kidney, liver, lung, salivary gland, spleen, testis and thymus. To determine if the endogenous *Lsm16* gene is expressed in adult BAT, WAT, and muscle (soleus and gastrocnemius), Northern blot analysis was conducted here using total RNA from these organs, as well as liver and brain to verify the previous results. The endogenous *Lsm16* transcript of approximately 3.3 kb was expressed in all tissues analyzed (Fig. 29). This confirms the expression of the gene throughout the adult mouse.

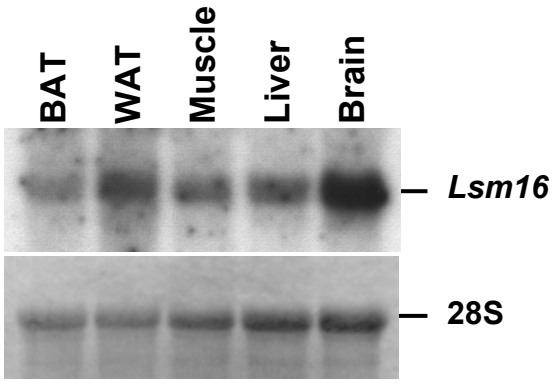


Figure 29. Northern blot analysis of total RNA from adult wild-type mouse tissues.
The endogenous *Lsm16* transcript was found at the expected size of approximately 3.3 kb in all tissues analyzed. The 28S rRNA band was used as loading control.

3.3.5. Analysis of hindleg muscle in the A008A01 mouse line

3.3.5.1. Lsm16/β-gal fusion protein is differentially expressed in leg muscles

Northern blot analysis had demonstrated that *Lsm16* is expressed in muscle. To determine the distribution of the expression in soleus and gastrocnemius muscles, cryostat sections of muscle of homozygous mutants at 6 months of age were stained with X-gal to detect the Lsm16/β-gal fusion protein (Fig. 30). In both muscles, punctate expression of the fusion protein was observed predominantly in the cytoplasm and on the periphery of the muscle fiber. Some nuclei appeared to contain staining. Interestingly, soleus muscle was stained in every muscle fiber, with approximately 4-10 foci of stain per fiber, whereas gastrocnemius muscle was stained in only a small proportion of the fibers, with only 1-3 foci per fiber. The main

functional difference between the soleus and gastrocnemius muscles is based on fiber type, soleus muscle being composed of a combination of type I and type II fibers, whereas gastrocnemius muscle is almost entirely type II fibers. Staining was not found exclusively in one type or the other. These results indicate that although the gene is expressed in gastrocnemius and soleus muscles to varying degrees, it is not associated with a specific fiber type.

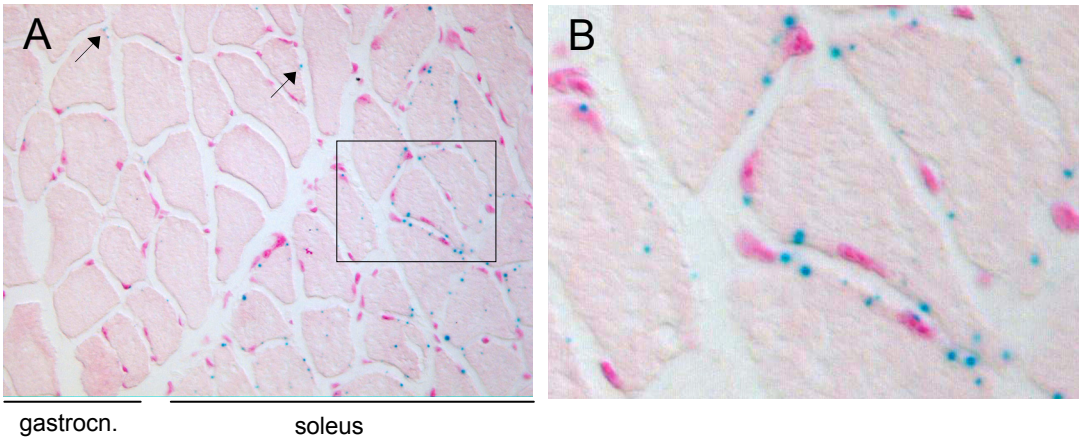


Figure 30. Detection of the fusion protein in muscle of a homozygous mutant.
(A) β-galactosidase activity of the fusion protein is found in each fiber of the soleus muscle, but only in a limited number of fibers of the gastrocnemius muscle (arrows indicate examples). (B) At higher magnification the staining in the soleus muscle can be seen predominantly as peripheral foci. 8 µm cryosection, counterstained with nuclear fast red to visualize nuclei.

3.3.5.2. No difference in muscle fiber-type distribution is observed between wild-type and homozygous mutant mice

A previous observation had suggested a switch in muscle fiber type had occurred in mutants, which might provide a link to their reduced performance in the Hanging-Wire Test. Fiber type distribution was examined on cryostat sections of soleus muscles of wild-type and homozygous mutant mice. As mentioned above, soleus muscle consists of a combination of type I and type II fibers. Type I fibers derive their energy from oxidative metabolism and are adapted for sustained contraction, whereas type II fibers, which derive their energy primarily from anaerobic glycolysis are adapted for rapid contractions of brief duration.

The fiber type was determined using myofibrillar ATPase (mATPase) staining. This method allows for an optical differentiation of fiber type. Fibers with high mATPase activity, i.e. type I, display dark blue staining, whereas fibers with low mATPase activity, i.e. type II, express light blue staining.

Quantification of the fibers indicated that the average difference between wild-type (Fig. 31A) and mutant (Fig. 31B) was less than 1% in either of the fiber types (Fig. 31C). In addition, the values do not vary from published values for the soleus from wild-type mice of 34 ± 2 % for type I and 66 ± 3 % for type II fibers (Totsuka *et al*., 2003). These results demonstrate that no switch in fiber type is occurring in the soleus muscle of mutant mice.

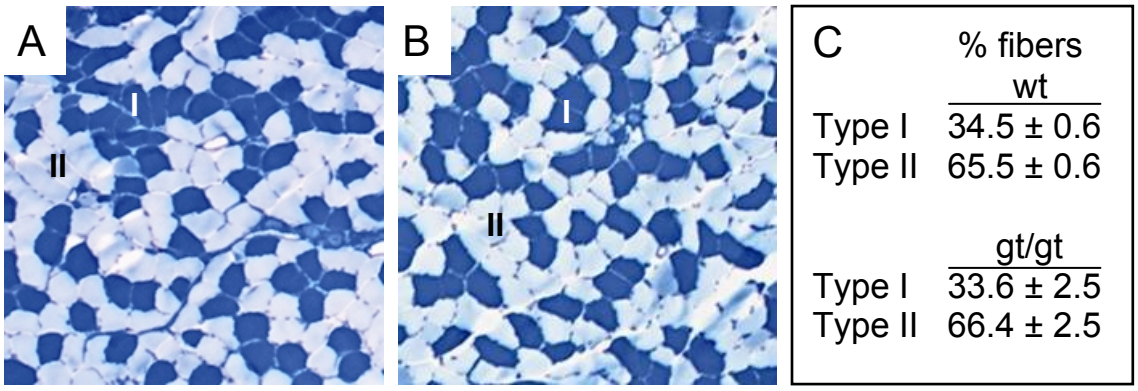


Figure 31. Sections of mATPase stained wild-type and mutant soleus muscle.
Soleus muscle from wild-type (A) and homozygous mutant (B) mice was cryosectioned at 10 µm and stained for mATPase activity to determine type I and type II fiber distribution. (C) Assessment of fiber types in wild-type and mutant mice. Data represents the mean ± s.d. in two animals within each group.

After comparing mATPase activity and Lsm16/β-gal fusion protein activity in serial sections of gastrocnemius and soleus muscles, it could be reconfirmed that no correlation between fiber type and fusion protein expression was found (data not shown).

3.3.5.3. Hindlegs of mutant mice are reduced in size

To determine if the difference in performance of mutants observed previously in the Hanging-Wire Test was muscle related, sections of the lower hindleg from 1 year old wild-type and homozygous mutant mice were compared for anatomical differences. In transverse sections, the hindlegs of the homozygous mutants (Fig. 32B) demonstrated a reduction in total size as compared to wild-type littermates

(Fig. 32A). This is most likely attributable to the general reduction in body size that has already been described. In addition, the individual muscles exhibited a proportional reduction in size in mutants. These results indicate that there is no obvious anatomical defect in hindlegs.

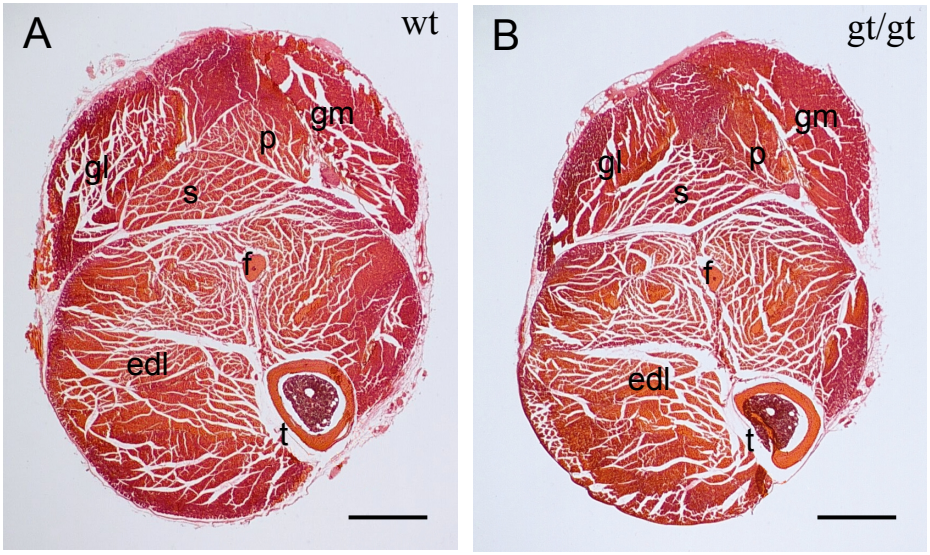


Figure 32. Comparative morphology of cross-sections through the lower leg.
In comparison to wild-type (A), the homozygous mutant (B) displays a reduction in the total size of the leg in cross-section, as well as the individual muscle groups. Hindlegs were fixed in Bouin's fixative, sectioned at 16 µm, and stained with hematoxylin and eosin. edl: extensor digitorum longus, f: fibula, gl: gastrocnemius lateral, gm: gastrocnemius medial, p: plantaris, s: soleus, t: tibia. Bar=1 mm

Taken together, no anatomical defects could be found in hindleg muscles of mutants. The smaller size of hindlegs in cross-sections reflects the reduction of size seen in the whole mouse. In addition, neither fiber-type switching nor a correlation between fiber type and expression of the fusion protein were detected in homozygous mutants. These results do not support the earlier observations suggesting a fiber-type switch and muscle weakness in homozygous mutants.

3.3.6. Lsm16/β-gal fusion protein is expressed in fat tissues

The greatest weight reduction in percent between wild-type and mutant mice was seen in white adipose tissue. To confirm the expression of the Lsm16/β-gal fusion protein in fat tissues, cryostat sections of WAT (Fig. 33A) and BAT (Fig. 33B) of homozygous mutants were stained with X-gal. Punctate staining was observed in most cells of both tissues, but with variable intensity and foci number. Similar to in muscle, staining was seen predominantly cytoplasmic.

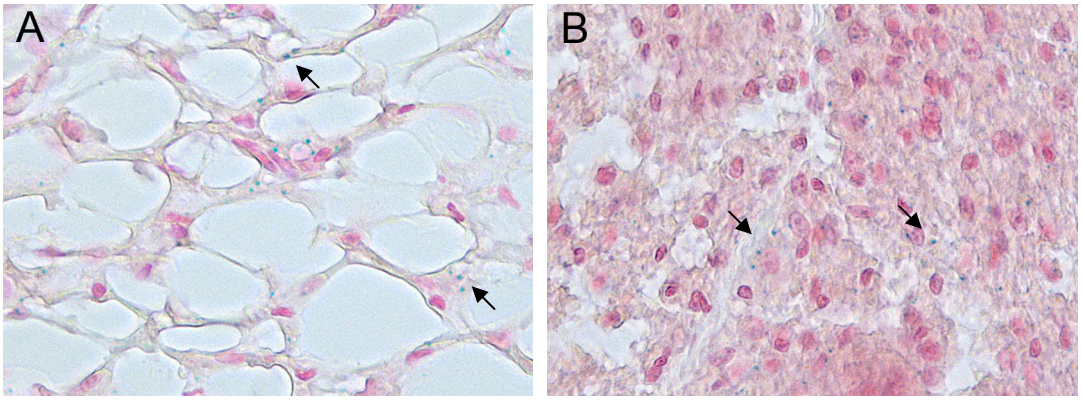


Figure 33. Detection of the β-gal fusion protein in white and brown adipose tissues.
The activity of the fusion protein was found as well defined foci throughout the WAT (A) and BAT (B) of homozygous mutants. Arrows indicate examples of more intensive staining. Cryosections of 10 µm, counterstained with nuclear fast red to visualize nuclei.

3.3.7. Altered expression of energy metabolism and adipocyte marker genes in mutants

The phenotype of homozygous mutant mice suggested energy metabolism and regulation might be affected by the mutation in the A008A01 mouse line. To explore the effects of the deficiency of *Lsm16* on the expression of marker genes characteristic for energy metabolism and regulation, as well as adipocyte specific genes, RT-PCR was performed. The goal of this analysis was to detect changes in expression of marker genes to explain the mouse phenotype.

Each of the tissues analyzed, i.e. BAT, WAT, muscle, liver and brain, plays a role in energy regulation and metabolism. While WAT is most importantly a source of energy in the form of lipids, BAT is a specialized thermogenic tissue responsible for metabolic heat production, or non-shivering thermogenesis. The liver is the site of entry to the body of all foods following digestion. It maintains circulating nutrient levels to supply other tissues by playing important roles in the metabolism of glucose, fat and amino acids. The brain, on the other hand, does not store any fuel reserves, but is one of the major energy consumers, as well as being at the center of energy homeostasis and body weight control. Muscle is another major energy consumer. It uses glucose from its own supply of stored glycogen, and other fuels supplied to it from other tissues by the blood. Hence, the genes tested by RT-PCR were characteristic for functions of adipose tissues and various aspects of energy metabolism and regulation (Tab. 7).

A first set of RT-PCR experiments was conducted using pools of reverse transcribed RNA from 2 wild-type and 2 homozygous mutant females, 16 months of age. Of the 21 genes analyzed (Tab. 7), only 7 showed a difference in expression between wild-type and homozygous mutant mice in at least one tissue (data not shown). Interestingly, 4 of the 7 genes, i.e. *adiponectin*, *resistin*, *adipsin* and *tumor necrosis factor α*, come from the same category, adipocyte derived hormones. Of the remaining three, *glycerol-3-phosphate acyltransferase* and *fatty acid synthase* both belong to the group of genes associated with triglyceride synthesis, and *sterol O-acyltransferase 1* is involved in cholesterol metabolism.

Transcription factor associated with adipogenesis
Peroxisome proliferator activated receptor gamma (*Pparγ*)

Adipose differentiation marker
Lipoprotein lipase (*Lpl*)

Adipocyte derived hormones
Adiponectin (*Adipoq*)
Leptin (*Lep*)
Resistin (*Retn*)
Adipsin (*Adn*)
Tumor necrosis factor alpha (*Tnfα*)

Mediator of lipolysis
Lipase, hormone sensitive (*Lipe*)

Transcription factor controlling triglyceride synthesis and cholesterol metabolism
Sterol regulatory element binding factor 1 (*Srebf1*)

Genes associated with triglyceride synthesis (lipogenesis)
Glycerol-3-phosphate acyltransferase, mitochondrial (*Gpam*)
Diacylglycerol O-acyltransferase 1 (*Dgat1*)
Diacylglycerol O-acyltransferase 2 (*Dgat2*)
Fatty acid synthase (*Fasn*)
ATP citrate lyase (*Acly*)

Genes associated with cholesterol metabolism
3-hydroxy-3-methylglutaryl-Coenzyme A synthase (*Hmgcs1*)
3-hydroxy-3-methylglutaryl-Coenzyme A reductase (*Hmgcr*)
Low density lipoprotein receptor (*Ldlr*)
Sterol O-acyltransferase 1 (*Soat1*)

Genes associated with insulin signaling
Insulin receptor (*Insr*)
Insulin receptor substrate 1 (*Irs1*)
Solute carrier family 2 member 4 (*Slc2a4*)

Table 7. Genes selected for RT-PCR analysis.
Genes having important roles in energy regulation and metabolism in adipocytes and other tissues were selected for analysis of their expression in BAT, WAT, muscle, liver and brain of mice in the A008A01 line. Genes are categorized by general function and listed by name, as well as symbol in parentheses. *Slc2a4* is more commonly known by its alias *Glut4*.

A second set of RT-PCR experiments was conducted using pools of reverse transcribed RNA from 3 wild-type and 4 homozygous mutant males, 5 months of age. Here, only the 7 genes showing expression differences mentioned above were analyzed. Similar results were obtained as with the first experiments. To better quantify the differences in expression between wild-type and homozygous mutant mice in the second set of experiments, densitometric analysis of the bands was performed. The wild-type expression values were set at 100% and the homozygous mutant expression values were normalized for loading using Rpl7 and calculated as percent of transcript quantity as compared to wild-type (See Materials and Methods 2.2.1.5.1 for a more detailed explanation of the procedure). Each PCR reaction was performed twice and analyzed. The average of the reactions was taken as the final result. Because each wild-type transcript quantity was set at 100%, the results in figures 34-36 do not reflect absolute transcript levels between tissues.

Among the adipocyte derived hormones, *adiponectin*, *resistin*, *adipsin* and *tumor necrosis factor α* all showed expression differences in at least 4 of the 5 tissues analyzed. It should be noted, although these proteins are referred to as adipocyte derived hormones, they are expressed in each of the tissues analyzed.

In homozygous mutants, the expression of *adiponectin* in brown and white adipose tissues was 92% and 80%, respectively, relative to the wild-type levels (100%). This represents slight decreases of 8% and 20%, respectively. Transcript levels were increased in muscle, brain and most strongly in liver by 28%, 121% and 217%, respectively (Fig. 34A). *Adipsin* expression was decreased in mutant WAT (35%), but greatly increased in BAT (181%) (Fig. 34B). Increases were also seen in muscle, liver and brain, 131%, 74% and 139%, respectively. In homozygous mutants, *resistin* exhibited elevated transcript levels in all tissues except WAT (Fig. 34C). The increase in BAT was 98%, in muscle 200%, in liver 270%, and in brain 62%. The 2% decrease in WAT is virtually equal to wild-type levels. *Tumor necrosis factor α* expression levels were increased in all 5 homozygous mutant tissues (Fig. 34D). The increase in BAT was 47%, in WAT 26%, in muscle 269%, in liver 270%, and in brain 171%.

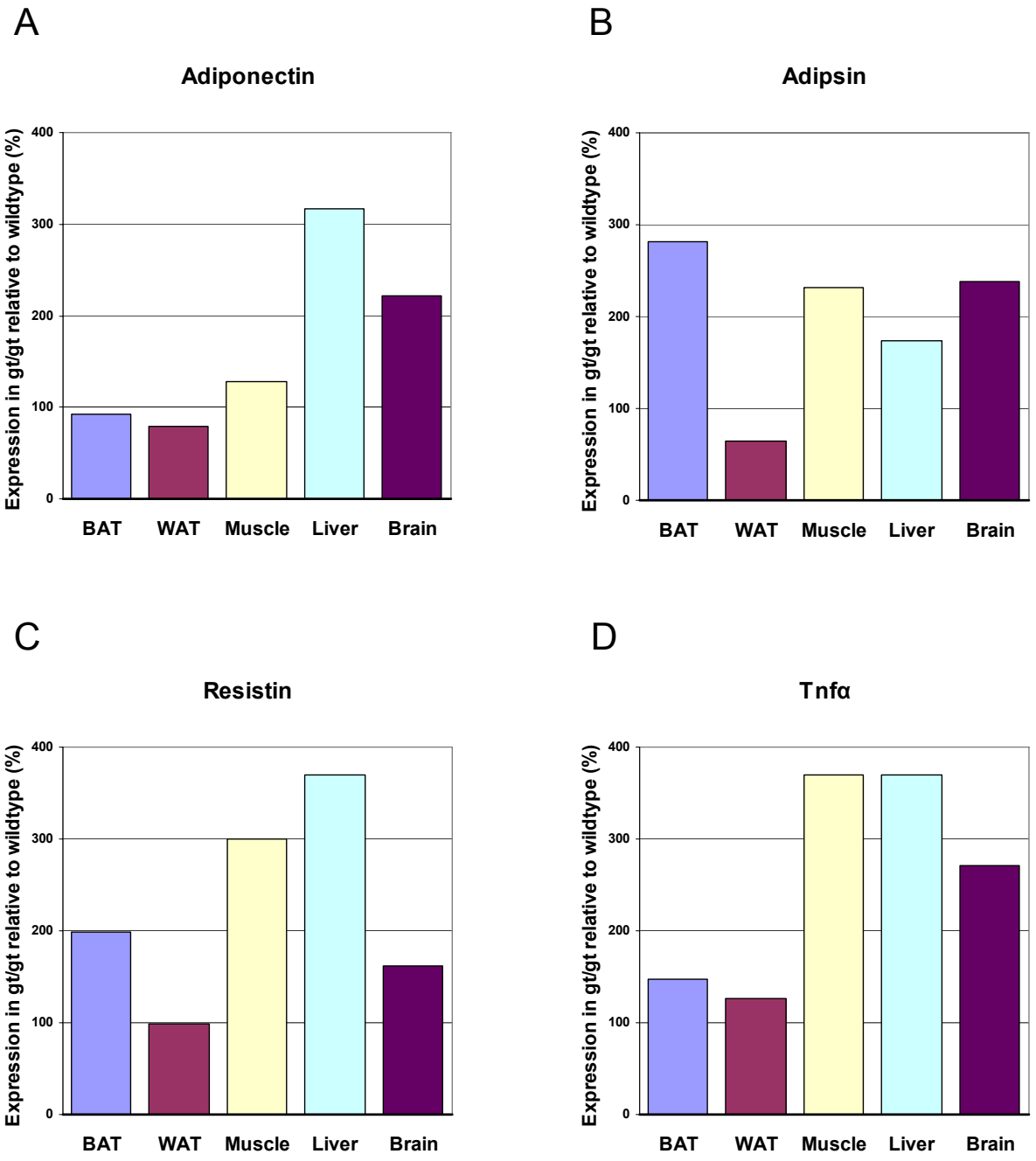


Figure 34. Adipocyte derived hormones: transcript levels in homozygous mutant mice relative to wild-type. Expression levels of *adiponectin* (A), *adipsin* (B), *resistin* (C) and *Tnfα* (D) in brown and white adipose tissue (BAT and WAT), muscle, liver and brain of homozygous mutant mice as compared to wild-type. Wild-type transcript levels (not shown) represent 100%.

Among the genes selected that are associated with triglyceride synthesis, only *glycerol-3-phosphate acyltransferase* (*Gpam*) and *fatty acid synthase* (*Fasn*) were found to have altered expression in at least one of the tissues tested. *Gpam* displayed slight increases in expression levels of 18% and 25% in muscle and brain, respectively (Fig. 35A). A greater increase of 48% was found in WAT. In BAT a 33% decrease in expression was seen, whereas no difference to wild-type was found in liver. In homozygous mutants, *Fasn* exhibited elevated transcript levels in all tissues except BAT (Fig. 35B). The increase in WAT was 25%, in muscle 105%, in liver 56%, and in brain 25%. In BAT, expression was reduced by 6%.

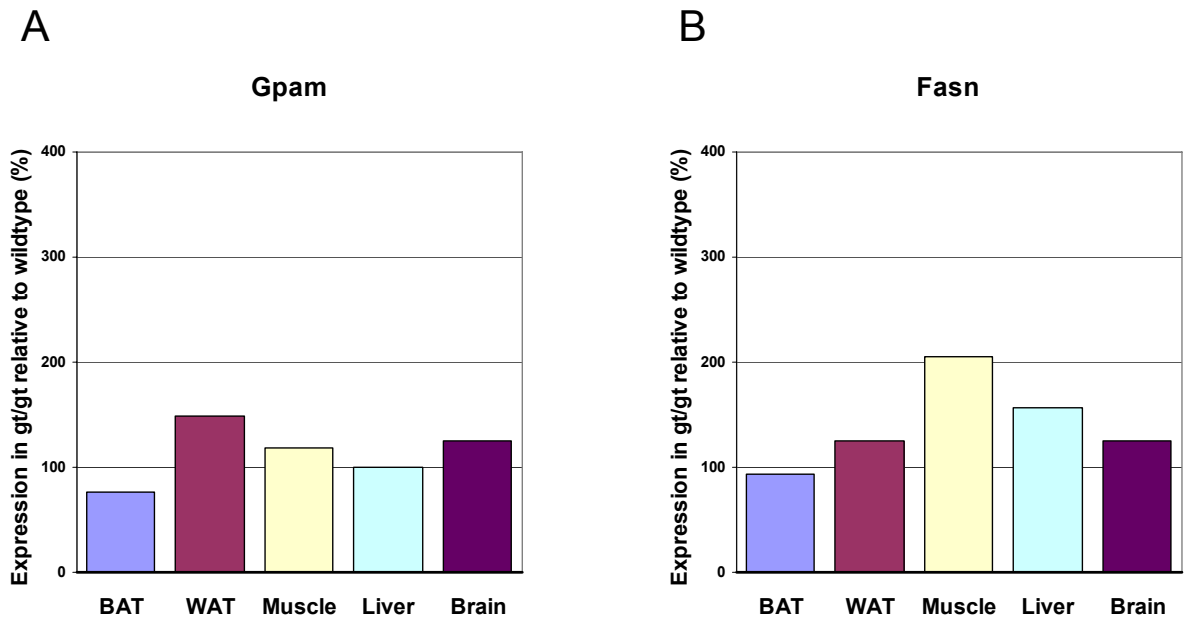


Figure 35. Genes associated with triglyceride synthesis: transcript levels in homozygous mutant mice relative to wild-type. Expression levels of *glycerol-3-phosphate acyltransferase* (*Gpam*) (A) and *fatty acid synthase* (*Fasn*) (B) in brown and white adipose tissue (BAT and WAT), muscle, liver and brain of homozygous mutant mice as compared to wild-type. Wild-type transcript levels (not shown) represent 100%.

Among the genes associated with cholesterol metabolism, *sterol O-acyltransferase 1* (*Soat1*) was the only one to show differences in expression levels. Expression of *Soat1* in homozygous mutants was reduced in both adipose tissues, BAT and WAT, 68% and 20%, respectively (Fig. 36). In muscle, liver and brain, expression was increased 61%, 161% and 68%, respectively above wild-type levels.

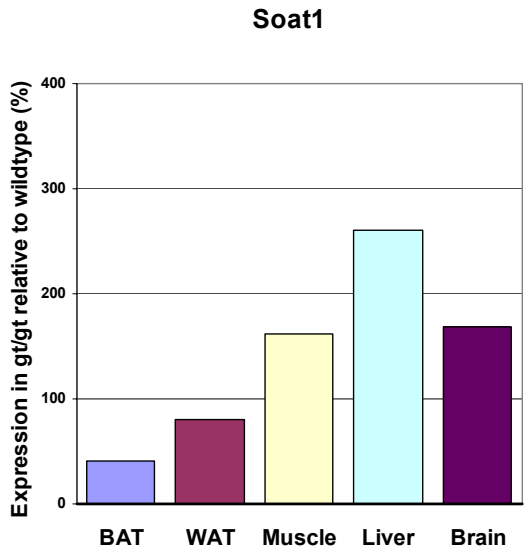


Figure 36. Sterol O-acyltransferase: transcript levels in homozygous mutant mice relative to wild-type. Expression levels of *sterol O-acyltransferase 1* (*Soat1*) in brown and white adipose tissue (BAT and WAT), muscle, liver and brain of homozygous mutant mice as compared to wild-type Wild-type transcript levels (not shown) represent 100%.

From the RT-PCR results it was found that 7 of the 21 genes analyzed exhibited differences in expression between wild-type and homozygous mutant mice in at least one of the tissues tested. Of the 7 genes, all showed increased expression in muscle, liver and brain, with the exception of *Gpam* in liver, which showed no change. Only in BAT and WAT did expression of the 7 genes vary above and below wild-type levels. Of the 5 tissues, WAT exhibited, in general, the least degree of change in homozygous mutants. With respect to the classification of the genes affected, 4 belonged to the category of adipocyte derived hormones, 2 to the category of genes associated with triglyceride synthesis and one to the category of genes associated with cholesterol metabolism. In each of these categories, though, there were genes whose expression was not affected.

4. Discussion

The focus of this work was to study gene function using mouse mutants generated by the gene trap technology. Since gene trapping leads to random mutations, known and unknown genes can be selected for analysis. Here, three mutant mouse lines were characterized concerning the *Arp3*, a KRAB-ZF, and the *Lsm16* genes. *Arp3* represents a known gene, KRAB-ZF is a member of a family of transcription factor genes, and the *Lsm16* gene is largely unknown, with essentially no previously characterized structural domains.

4.1. The *Arp3* mutant mouse line

The mutant mouse line A009F03 contains the gene trap vector integrated in intron 1 of the *Arp3* gene. This results in a null mutation, as the protein sequence encoded by exon 1 contributes only 14 amino acids to the deduced fusion protein and exons 2-12 are lost in the mutant.

Northern blot analysis has shown that *Arp3* is probably expressed in all organs, including brain, heart, intestine, kidney, liver, lung, salivary glands, skeletal muscle, spleen, testis and thymus. The ubiquitous expression underscores the known role of Arp3 in cells throughout the organism.

In heterozygous mutants the level of *Arp3* transcripts was reduced by approximately 50%. This decline also applied to the amount of protein. This observation confirms that the gene trap allele leads to loss of Arp3 protein and homozygous mutant mice are likely to represent an Arp3 null mutation. As heterozygous Arp3 mutant mice are viable and apparently not significantly different from wild-type mice, the dosage of Arp3 protein in cells does not seem to be critical. In contrast, mice homozygous for the gene trap integration are not found at birth, indicating that the mutation is recessive and causes prenatal lethality. Indeed, homozygous embryos can only be detected until the blastocyst stage at approximately E3.5. No morphological differences were observed between wild-type, heterozygous and homozygous mutant embryos at the stages of morula (E2.5), compact morula (E3.0), early blastocyst (E3.25) and almost fully expanded blastocyst (E3.5). Isolation of E3.5 embryos from heterozygous matings yielded all three genotypes. However, after one night of cultivation all homozygous mutant embryos failed to continue development,

while most of the wild-type and heterozygous blastocysts progressed through further development normally. From this, it can be concluded that Arp3 null-mutant embryos die during the blastocyst stage. The early lethality clearly indicates the absolute requirement of Arp3 protein for preimplantation development.

The necessity of Arp3 for the cell is also evident by the inability to obtain homozygous mutant ES-cells from heterozygous ES-cells following selection in high concentrations of geneticin (Freese, 2003).

Arp3 is known as one of the seven subunits in the Arp2/3 complex, which is a key regulator of actin filament polymerization (Mullins *et al*., 1998; Pollard and Beltzner, 2002). It is therefore reasonable to believe that the absence of a functional Arp2/3 complex and impairment of actin filament nucleation involving the Arp2/3 complex causes embryonic lethality. It has been demonstrated in *C. elegans* that disruption of Arp3 expression by RNAi leads to developmental arrest during morphogenesis with 100% lethality (Sawa *et al*., 2003). Here, hypodermal cell migration is inhibited during ventral enclosure owing to the reduction of filamentous actin formation. In *Drosophila*, a loss-of-function mutation for Arp3 is lethal in late stage egg chambers (Hudson and Cooley, 2002). Here, the failure of actin polymerization leads to defective expansion of ring canals, which affects the transfer of cytoplasm from germline nurse cells to the oocyte.

During preimplantation development of the mouse, the Arp2/3 complex is required for regulating actin filament polymerization in E-cadherin mediated adherens junctions (Kovacs *et al*., 2002; Verma *et al*., 2004). The most simplified structure of the adherens junction consists of the transmembrane protein E-cadherin, which is responsible for cell-cell adhesion, β-catenin, which binds to the cytoplasmic domain of E-cadherin, and α-catenin, which binds to β-catenin and connects this adhesion complex with the actin filament network (Nagafuchi and Takeichi, 1989; Ozawa *et al*., 1989). Adherens junctions first appear in the mouse embryo at the 8-cell morula stage (E2.5). There, they are responsible for establishing extensive cell-cell contacts between the undifferentiated blastomeres, which leads to compaction of the morula (Johnson *et al*., 1986). During the next two rounds of division, the outer cells of the compact morula become a fluid-transporting epithelium, the trophectoderm, which is

responsible for the formation of the blastocoel, the fluid-filled cavity of the blastocyst. The inner cells of the morula become the inner cell mass (ICM) of the blastocyst.

Null mutants of the main adherens junction proteins, E-cadherin, α-catenin and β-catenin, exhibit early embryonic lethal phenotypes (Larue *et al*., 1994; Riethmacher *et al*., 1995; Torres *et al*., 1997; Haegel *et al*., 1995). Although all null embryos undergo compaction due to residual maternal protein, in the case of E-cadherin and α-catenin, the null embryos are not capable of forming a complete trophectodermal epithelium and blastocoelic cavity. The authors propose that this results from the inability of the cadherin-catenin complexes to interact with actin filaments and thereby achieve a stable and functional epithelial junction. β-catenin null embryos survive to early gastrulation, most likely due to functional compensation by plakoglobin (Haegel *et al*., 1995).

Characterization of the distribution of Arp3 protein in preimplantation embryos by immunohistochemistry demonstrated a ubiquitous expression of the protein throughout the embryo. This was not only true for wild-type and heterozygous embryos between the stages E2.0 and E3.5, but also for homozygous mutants. Because these results were obtained using two different Arp3 specific antibodies directed at different epitopes, and because the negative controls, without the anti-Arp3 antibody, exhibited no staining, it can be concluded that the antibodies and staining procedure were not at fault. The most plausible explanation is the presence of maternal protein. This conclusion is also supported by the 50% reduction of Arp3 protein in heterozygous organs i.e. mutant alleles are incapable of producing protein, which would indicate that the homozygous mutant embryos would be incapable of expressing Arp3 protein encoded by their own genome.

In general, it is known that oocytes contain not only numerous maternal proteins, but that gene expression during early embryonic development until zygotic gene activation (ZGA) is regulated mainly by translational activation of maternally derived mRNAs. In the mouse, ZGA occurs at the two-cell stage, E1.5 (Flach *et al*., 1982). By this time, maternal mRNA degradation has been triggered and is ~90% completed.

Maternal protein synthesis continues though into the eight-cell morula stage, E2.5 (Piko and Clegg, 1982; Paynton *et al*., 1988; Schultz, 1993). The presence of maternal Arp3 protein at stage E3.5 would imply that maternal Arp3 synthesis continues beyond stage E2.5, or that the Arp3 protein or the Arp2/3 complex remains stable up to stage E3.5.

In the Arp3 gene trap line, normal morphology and survival of mutant embryos is seen up to stage E3.5. This is most likely accountable to maternal component. Hence, the lethal phenotype is first obvious after the concentration of maternal Arp3 protein and thus the concentration of Arp2/3 complex has decreased below a critical level for supporting actin filament polymerization. It is further proposed that this would lead to instability in adherens junctions and subsequent loss of the ability of the trophectoderm to further support the development of the embryo. This is visible by the growth arrest and deterioration of homozygous mutant embryos after stage E3.5. An additional possibility that could affect the development of the blastocyst entails endocytosis. Actin filaments nucleated by the Arp2/3 complex are involved in clathrin-mediated endocytosis (Engqvist-Goldstein and Drubin, 2003; Merrifield, 2004; Yarar *et al*., 2005). In mammalian cells this has a vital role in such cellular processes as nutrient uptake and cell surface receptor down-regulation. This has yet to be studied in mouse embryonic development though.

In summary, the analysis of the A009F03 mouse line has demonstrated that the Arp3 protein is distributed ubiquitously throughout preimplantation embryos, at least during stages E2.0 to E3.5, and this protein appears to be provided maternally. Although embryos are morphologically normal up to stage E3.5, homozygous embryos fail to develop beyond this stage. The lethality could be caused by the instability of adherens junctions leading to defective trophectoderm. Survival to stage E3.5 is probably the result of maternal protein.

Due to the Arp3 mutation leading to an embryonic lethal phenotype in homozygous mutants and no obvious phenotype in heterozygous mice, further analysis of this line will be difficult. Although the most likely cause of lethality is the loss of functional adherens junctions, and possibly disrupted endocytosis, this would still need to be verified. The simplest and most straight forward way of studying Arp3 and the Arp2/3

complex is via cell culture. It is therefore unfortunate that Arp3 homozygous mutant ES-cells could not be generated from the gene trap line.

4.2. The KRAB-ZF mutant mouse line

The mutant mouse line A20010 contains the gene trap vector integrated in a novel krüppel-associated box zinc-finger (KRAB-ZF) gene located on chromosome 7. The fusion protein encoded by the mutant allele is missing the zinc-finger motifs and nuclear localization signal, but still possesses the KRAB domain. It is therefore uncertain if this results in a null mutation. Although the fusion protein would not be able to fulfill its putative role as a transcriptional repressor, the KRAB domain might bind TIF1β or other proteins in the cytoplasm. A new function for the fusion protein seems unlikely, though, since no obvious phenotype has been detected in mutant mice.

Expression of the endogenous KRAB-ZF transcripts and KRAB/β-galactosidase fusion protein illustrates a very broad spatio-temporal expression pattern for the gene. The gene is expressed weakly, though, which is apparent by the failure to detect the fusion protein by X-gal staining, and the KRAB-ZF transcripts by Northern blotting. Furthermore, the transcripts could first be detected with 35-40 PCR cycles. Endogenous transcripts of the KRAB-ZF gene are found as a full-length and an alternatively-spliced variant. The loss of both of these transcripts in homozygous mutants indicates that mutagenesis by the gene trap vector is successful and that the PCR fragments representing these transcripts are specific for the trapped gene. It still needs to be verified that the splice variant encodes a protein as predicted. At present, the protein would consist of only the KRAB domain, making it similar to the endogenous portion of the KRAB/β-gal fusion protein, which also contains only the KRAB domain. The mutation in the full-length transcript should still lead to loss of function though.

Since expression of the KRAB-ZF gene does not appear to be specific to any developmental stage, tissue or cell type, and known KRAB-ZF proteins are involved in a broad range of biological process, it is not possible to determine the function of the KRAB-ZF trapped in the A20010 mouse line at this time. Furthermore, no observable phenotype is present in mutants.

Initially, a hypopigmentation phenotype was seen in A20010 mutants. This was evident by a dramatic decrease or complete loss of pigment in eyes, hair follicles and hair shafts of KRAB-ZF mutants. Therefore, analysis focused on the characterization of pigmentation in these tissues. It could be proven, however, that the phenotype was a result of mutant alleles of the *pink-eyed dilution* (*p*) and *tyrosinase* (*Tyr*) genes, in the case of *tyrosinase*, the chinchilla allele (Tyr^{ch}). These alleles came from the R1 ES-cells used to generate the A20010 mutant mouse line, and cosegregated with the KRAB-ZF mutant allele. R1 is a cell line derived from the inner cell mass of blastocysts isolated from crosses between 129/SvJ-Tyr^{ch} p/+^{Tyr} +^{p} and 129/Sv-+^{Tyr} +^{p} mice (Nagy *et al*., 1993; Simpson *et al*., 1997; Threadgill *et al*., 1997) The cells are therefore heterozygous for the *p* and Tyr^{ch} mutations.

Pink-eyed dilution is an integral, melanosomal membrane protein important for the proper function of melanosomes (Rosemblat *et al*., 1994). Research on p-deficient cell lines has indicated several possible functions of the p protein, including stabilization of the tyrosinase containing melanosomal protein complex (Lamoreux *et al*., 1995), regulation of the pH in melanosomes (Puri *et al*., 2000), and processing and transport of tyrosinase (Chen *et al*., 2002). Mutations at this locus have been characterized by a reduction of melanin synthesis in the skin, hair and eyes of mice and humans (Silvers, 1979; Witkop *et al*., 1972; King *et al*., 1985). This results in the disorder oculocutaneous albinism type II (OCA2) (Rinchik *et al*., 1993). In humans, OCA2 is the most common form of albinism worldwide (Ramsay *et al.*, 1992; Durham-Pierre *et al*., 1994). R1 ES-cells carry the *p* mutation from the 129 genetic background. This is a low activity mutation, not a null mutation, believed to cause a reduction of *p* transcript levels (Rinchik *et al*., 1993). In A20010 mice, genotyping the *p* gene proved that the mutant allele was present in KRAB-ZF mutant mice. In addition, reduced levels of *p* transcripts were found in homozygous mutant embryos. These results verify the involvement of the *p* mutation in the hypopigmentation phenotype of A20010 mutants. Since *p* mutations cause a stronger reduction in eumelanin than pheomelanin (Rinchik *et al*., 1993), this would explain the appearance of a switch from eumelanin to pheomelanin in the hypopigmented tissues.

Tyrosinase is also an integral, melanosomal membrane protein and is the rate-limiting enzyme in melanin biosynthesis. Its primary function is catalyzing the hydroxylation of tyrosine to dopaquinone (Cooksey *et al*., 1997). In mice and humans, oculocutaneous albinism type 1 (OCA1), the second most common form of albinism worldwide, is a disorder associated with mutations that lead to the production of inactivate tyrosinase, and thus a reduction in the amount of melanin (reviewed in Oetting, 2000). The chinchilla allele encodes a tyrosinase with a point mutation at position 464 (Ala464Thr). This leads to an increased susceptibility of the protein to proteolytic cleavage and a 3-fold reduction in activity (reviewed in Lamoreux *et al*., 2001). In A20010 mice, sequencing results confirmed the point mutation in KRAB-ZF homozygous mutant mice, hence verifying the participation of the *Tyr*^{ch} mutation in the hypopigmentation phenotype of A20010 mutants. The chinchilla mutation does not explain the RT-PCR results showing two *tyrosinase* splice variants, i.e. full-length and Δ3, in all mice. In wild-type mice, the Δ3 splice variant is known to appear in reduced amounts together with the full-length transcript, but to encode an inactive enzyme (Terao *et al*., 1989; Porter and Mintz, 1991). In A20010 mutants, there appears to be a switch in *tyrosinase* expression from the full-length transcript to the Δ3 transcript, thus from an active to an inactive protein. It is not known if this might be an additional effect of the chinchilla, p or KRAB-ZF mutations. Such a switch has been found to be associated with pigmentary decrease in melanomas (Le Fur *et al*., 1997).

It is likely that the *p* mutation alone is responsible for the intermediate pigmentation seen in beige and light gray mice, if the chinchilla allele is lost due to recombination. Wild-type pigmentation in KRAB-ZF homozygous mutants probably results from the loss of both *p* and *Tyr*^{ch} mutations by recombination.

In summary, although the KRAB-ZF gene appears to be expressed throughout the mouse at all stages of life, and as two splice variants, the mutation of the gene leads to no obvious phenotype. The expression of the fusion protein was not observed in all cells, though. Keeping in mind that KRAB-ZF proteins are known as transcriptional repressors, the spatio-temporal expression pattern of the trapped gene does not provide enough information to determine a target gene or process that could be affected. It must be assumed that the KRAB-ZF protein represses a gene or genes

that have a subtle role in the mouse or that the loss of the protein is being compensated for by a redundant gene.

Since the sequence or location of a gene does not need to be known in order to mutate it using the gene trap approach, this can lead to problems. It was the unfortunate situation that vector integration in the A20010 line was on chromosome 7 adjacent to the *p* and *Tyr* mutations, and that this was known only later. Following elucidation of the hypopigmentation phenotype and detection of no additional phenotype, further analysis with the mouse line was discontinued. A small number of mice are being kept and allowed to age to look for a later appearing phenotype. If the analysis of the A20010 mouse line were to be continued, most important for the future work would be finding a perspective gene or genes that act as targets for the KRAB-ZF protein. Here, it would be informative to use microarray analysis.

4.3. The *Lsm16* mutant mouse line

The mutant mouse line A008A01 contains the gene trap vector integrated into a gene, which until recently was a novel gene known only to have a YjeF-N domain. In this thesis, analysis of the amino acid sequence of the trapped gene revealed, in addition to the C-terminal YjeF-N domain, an N-terminal Lsm16-N domain and a central FDF domain. The presence of the Lsm16-N domain has lead to the classification of this protein as the Lsm16 ortholog in mouse. By comparing protein sequences, it was found that Lsm16 proteins are evolutionarily highly conserved, with the mouse protein exhibiting over 95% identity to its ortholog in other mammals.

Lsm16 belongs to the family of like-Sm (Lsm) proteins. Sm and Lsm proteins share a common domain, the Sm-fold. The Lsm16-N domain is homologous to the Sm-fold, which has been characterized in Lsm1-8 and the members of the Sm family (Khusial *et al*., 2005). Lsm1-8 and Sm proteins are involved in mRNA processing by forming heterohexameric or heteroheptameric rings in which the Sm-fold of each protein is positioned at the lumen of the ring. The folds are necessary for ring formation and function collectively to bind RNA. Sm proteins are known to be components of small nuclear ribonucleoproteins (snRNPs) in the nucleus. These combine to form splicesomes that are involved in pre-mRNA splicing (Seraphin, 1995; Hermann *et al*., 1995). Similar to the Sm proteins, Lsm proteins are involved in general RNA

maturation in the nucleus, but in addition, they have been found to be involved in degradation of mRNA in the cytoplasm (He and Parker, 2000).

Recent work by Fenger-Gron and colleagues (2005) demonstrated that hEdc3, the human Lsm16 ortholog, coimmunoprecipitated with Rck/p54, hDcp1 and TTP. These are three factors that play important roles in enhancing mRNA decapping as part of the mRNA decay pathway. hEdc3 was also found to localize to processing bodies (P-bodies) that were visualized as cytoplasmic foci in HeLa cells. P-bodies represent multiprotein complexes responsible for mRNA decay in the cytoplasm (Cougot *et al*., 2004).

Similar results were found in *Saccharomyces cerevisiae* by Kshirsagar and Parker (2004). They reported that Edc3p, the yeast Lsm16 ortholog, had extensive two-hybrid interactions with several proteins involved in mRNA decapping and 5´-3´ degradation. In addition, although Edc3p was not required for decapping, it could stimulate decapping, and it localized to P-bodies. The authors postulate that Edc3p belongs to a class of decapping factors, which, although they can affect the decapping process, are not normally rate limiting for decapping. The role of Edc3p may lie in stabilizing or affecting the function of the decapping enzymes Dcp1 and Dcp2.

The results from research with human cells and yeast lead to the presumption that the Lsm16 protein would carryout the same function of enhancing mRNA decapping in the mouse. The protein would be one subunit in a heptameric ring with the Lsm16-N domains of each subunit forming Sm-folds at the lumen of the ring and binding mRNA for decapping. Although the FDF domain is yet uncharacterized, sequence and structural aspects of the domain imply that it interacts with RNA or highly charged peptides that are commonly found in ribonucleoprotein complexes (Anantharaman and Aravind, 2004; Albrecht and Lengauer, 2004). The YjeF-N domain is still uncharacterized and no function has been inferred.

In the A008A01 mutant mouse, approximately one-half of the endogenous protein sequence (i.e. amino acids 274-508) is replaced by β-galactosidase and neomycin phosphotransferase in the Lsm16/β-galactosidase fusion protein. The amino acid

sequences of the Sm-fold, i.e. the Lsm16-N domain, and three-fourths of the FDF domain are retained. This might be sufficient for the fusion protein to be partially functional or even lead to a new function in the cell. Ideally, the gene trap vector should have integrated into intron 2, which would have disrupted the Lsm16-N domain encoded by exons 2 and 3.

Detection of the *Lsm16* endogenous transcript by Northern blot analysis demonstrates that the gene is expressed in most adult tissues, the heart being an exception (Meyer, 2003). Of the tissues analyzed in this thesis, *Lsm16* transcript levels appear highest in WAT and brain, intermediate in liver and skeletal muscle (soleus and gastrocnemius) and lowest in BAT. Although the higher expression in WAT suggests that the gene has an increased function there, thus the mutation exhibits a more dramatic effect in WAT reduction, the same should then also be true of the brain. The brain, though, has the least weight difference between wild-type and homozygous mutant mice. Hence the level of expression of the gene in a tissue does not appear to directly correlate with the weight reduction in that tissue.

Detection of the Lsm16/β-gal fusion protein shows that the gene is expressed in all embryonic stages, E7.0 to E17.5, and throughout the embryo, with exception of the liver. In adult mice, expression is seen in most tissues, with exception of the heart (Meyer, 2003). There is no clear explanation why expression of the fusion protein is found in the adult liver, but not the embryonic liver, whereas expression is found in the embryonic heart, but not in the adult heart, especially since both of these organs show reduced weight in adult homozygous mutants. The general expression pattern of *Lsm16* throughout the embryo and adult mouse imply that the protein is involved in a process common to most cells of the organism. Since the fusion protein is expressed in all embryonic stages, but no phenotype has been detected in embryos, this indicates that the gene is not essential. If the protein is indeed involved in mRNA decapping, this would support the suggestion that it only acts as an enhancer. Localization of the expression of the Lsm16/β-gal fusion protein was examined in WAT, BAT and skeletal muscle. Staining is observed as punctate foci predominately in the cytoplasm in all three tissues. In WAT, BAT and soleus muscle this appears to be in all cells, whereas in gastrocnemius muscle staining is exhibited in very few of the muscle fibers, and with a reduced number of foci. The staining difference in

muscles does not correlate with fiber type. Since the mouse Lsm16 protein is homologous to the human hEdc3 and yeast Edc3p proteins, which are known to be found in P-bodies (Fenger-Gron *et al*., 2005; Kshirsagar and Parker 2004), this would suggest that the murine endogenous Lsm16 protein would also localize to P-bodies. This could explain the foci of β-gal staining. Although the cellular process involved in localizing these proteins to P-bodies is not known, if the expression of the Lsm16/β-gal fusion protein corresponds to P-bodies, then the sequence necessary for this is not missing in the fusion protein. Thus, the Lsm16-N domain and part of the FDF domain would to be sufficient for this localization. Of those foci found in the nucleus in mutant mice, the expression could represent an involvement of the protein in RNA processing in the nucleus, though a nuclear localization signal has not been found in the protein sequence. For more accurate localization, expression of the endogenous protein should be analyzed in cultured cells. With respect to the β-gal staining difference seen in muscle, if the Lsm16 protein is involved in mRNA decapping, the difference in expression of the fusion protein between soleus muscle and gastrocnemius muscle could simply reflect a higher rate of decapping or mRNA turnover in soleus muscle.

Further analysis of hindleg muscles failed to find anatomical differences or a fiber type switch between homozygous mutant and wild-type littermates. Hence, these results do not support the earlier suggestions that homozygous mutants exhibit muscle weakness and a fiber type switch.

Homozygous mutants are smaller in size and have a reduced weight as compared to their wild-type littermates. This phenotype first becomes obvious around 3 weeks after birth, with the average weight difference increasing with age. Although the lower weight in mutants is reflected in a reduction in weight of all tissues, each tissue is affected to a different extent. Here, white adipose tissue demonstrates the greatest difference, whereas the brain the least. Due to the phenotype in homozygous mutants it was presumed that energy metabolism was affected, as well as possibly a defect in adipogenesis, adipocyte function or survival. Of 21 marker genes for adipocytes and energy metabolism, whose expression was analyzed in BAT, WAT, muscle, liver and brain, 7 displayed noticeably altered transcript levels in homozygous mutants. These were the adipocyte derived hormones *adiponectin*,

resistin, *adipsin* and *tumor necrosis factor α*, the triglyceride synthesis associated genes *glycerol-3-phosphate acyltransferase* and *fatty acid synthase*, and the cholesterol metabolism associated gene *sterol O-acyltransferase 1*.

In homozygous mutants, the expression of *adiponectin* was increased in muscle, brain and most strongly in liver, whereas in BAT and WAT it was slightly reduced. Adiponectin is involved in controlling glucose and lipid metabolism. It enhances insulin sensitivity in liver and muscle, which leads to suppressed glucose production in liver and increased glucose uptake and glycogen accumulation in muscle (reviewed in Wong *et al*., 2004). In adipose tissues, adiponectin promotes adipogenesis and increases lipid content in adipocytes (Fu *et al*., 2005). The reduction of *adiponectin* expression in BAT and WAT suggests a reduction in the number of adipocytes, or differentiated adipocytes in fat tissues, together with a decrease in lipid content in cells of mutants. Although the function of adiponectin in the brain is yet unknown, it may have an effect on energy expenditure being mediated by the hypothalamus (Qi *et al*., 2004).

Adipsin expression was increased in BAT, muscle, liver and brain, but decreased in mutant WAT. Adipsin is a serine protease involved in stimulating lipogenesis and triglyceride synthesis. Its expression is induced upon differentiation of preadipocytes (Cook *et al*., 1985). The decrease in expression in WAT would suggest that, as with adiponectin, there may be a deficit in adipocyte differentiation and lipid content. The increase of almost 2-fold in expression in BAT though, would infer that adipsin's role in BAT is different than that in WAT. Although a further role of adipsin in metabolism is unknown, it has been identified as the same protein as complement factor D, where it has a proteolytic function in the alternative pathway of the complement system of immunity (Rosen *et al*., 1989).

As with *adipsin*, *resistin* exhibited elevated transcript levels in all organs except WAT, which was equal to the wild-type level. Resistin is known to be upregulated during adipogenesis (Steppan *et al*., 2001). The near wild-type expression level in mutant WAT implies that resistin is not directly affecting WAT in mutant mice. In adipocytes, as well as in other tissues, resistin has been found to enhance insulin resistance. In the liver, this leads to increased glucose production by glycogenolysis and

gluconeogenesis (Rangwala *et al*., 2004). Resistin displays the opposite effect on insulin sensitivity as that of adiponectin.

Tumor necrosis factor α expression levels were increased in all 5 homozygous mutant organs. Tnfα has multiple roles in lipid metabolism (reviewed in Yu and Ginsberg, 2005). In white adipose tissue, Tnfα causes reductions in the expression of genes involved in adipogenesis and lipogenesis. The increase in Tnfα may therefore be involved in the reductions of *adiponectin* and *adipsin* expression levels in WAT, thus leading to the total reduction in WAT. In addition, Tnfα also promotes lipolysis in adipocytes, leading to an elevation of plasma fatty acid levels, a hallmark of insulin resistance. In liver, Tnfα increases the expression of genes involved in *de novo* fatty acid synthesis e.g. *Fasn*, while decreasing expression of those involved in fatty acid oxidation. In various tissues, Tnfα has also been found to have a negative regulatory effect on glucose import.

Interestingly, leptin was the only adipocyte-derived hormone that did not display altered expression levels in homozygous mutants. Leptin, which acts as a signal of energy stores, improves insulin sensitivity, similar to adiponectin, and is responsible for inhibiting food intake and accelerating energy metabolism (Shulman, 2000). There is therefore no clear explanation why the other hormones would be affected, but leptin not, especially in adipose tissue.

Adipocyte derived hormones are involved in a myriad of processes with varying effects in different tissues. These have a combined effect on energy homeostasis throughout the mouse. From the known functions of the hormones alone, it is difficult to find a direct correlation between altered expression levels and the mutant mouse phenotype. The reductions of *adiponectin* and *adipsin*, together with the increase of *Tnfα* in WAT suggest a defect in adipogenesis or lipid content, though the other marker genes for adipogenesis and adipose differentiation, *Pparγ* and *Lpl*, are not affected. Similarly, the increase of the four hormones in liver would suggest that insulin sensitivity is affected, but the three genes involved in insulin signaling, *Insr*, *Irs1* and *Slc2a4*, show no altered expression.

Compared to the adipocyte-derived hormones, the triglyceride synthesis associated genes *glycerol-3-phosphate acyltransferase* (*Gpam*) and *fatty acid synthase* (*Fasn*) both exhibited less dramatic expression differences between mutant and wild-type mice. *Gpam* displayed slightly increased expression levels in WAT, muscle and brain, though the highest was only a 48% increase as seen in WAT. In BAT a decrease in expression was observed, whereas no difference to wild-type was found in liver. Gpam is an integral protein of the mitochondrial outer membrane where it catalyzes the initial and committed step in glycerolipid biosynthesis (Coleman *et al*., 2000). In addition, Gpam plays important roles in the synthesis of triacylglycerol, the fatty acid content of triacylglycerol, and the positioning of specific fatty acids in phospholipids (Hammond *et al*., 2002). From its role in lipogenesis it would be anticipated that *Gpam* levels be reduced in the mutant mice. This is only observed in BAT though.

In homozygous mutants, *Fasn* exhibited elevated transcript levels in all organs except BAT. Fasn plays an important role in energy homeostasis by synthesizing lipids for storage. The function of Fasn consists of catalyzing the synthesis of the saturated fatty acids myristate, palmitate, and stearate by using the substrates acetyl-CoA, malonyl-CoA, and NADPH (Chirala *et al*., 2003; Wakil, 1989). As with Gpam, the importance of Fasn for lipogenesis would suggest that the expression of the gene would be reduced in mutant mice. This was also only observed in BAT.

Expression of *sterol O-acyltransferase 1* (*Soat1*) in homozygous mutants was reduced in both adipose tissues, BAT and WAT, whereas in muscle, liver and brain expression was elevated. Soat1 is an integral membrane protein of the endoplasmic reticulum where it plays a role in lipoprotein assembly and dietary cholesterol absorption. This consists of converting unesterified cholesterol, derived from the uptake of modified lipoproteins, into esterified cholesterol in response to cholesterol abundance inside the cell (Brown *et al*., 1980). These results do not allow for a clear interpretation since the upstream genes involved in controlling, synthesizing and uptake of unesterified cholesterol, i.e. *Srebf1*, *HMGCS1*, *HMGCR*, and *Ldlr* do not show changes in transcript levels. In addition, blood serum levels of cholesterol were not significantly different between wild-type and homozygous mutant mice.

As with the adipocyte derived hormones, the results from *Gpam*, *Fasn* and *Soat1* do not allow for a direct association of altered gene expression to mouse phenotype. In the case of *Gpam* and *Fasn* this is even more difficult due to the results being the converse of what would be expected. This demonstrates the intricacy of energy metabolism.

Energy metabolism is a complex set of processes controlled by cross-talk of adiposity-related signals and nutrient-related signals. In simplest terms, nutrient availability in the body affects both nutrient-related signals and body fat mass. The fat mass then affects adiposity-related signals. Both types of signals are processed by the brain which responds by adjusting hepatic glucose production and energy balance, i.e. food intake and energy expenditure. These adjustments then influence nutrient availability in the body, closing the circular pathway of the regulation of energy metabolism (Lazar, 2005). This indicates that the reduction in size of the mouse in general, is a complex process with numerous genes being involved, which have varied effects in different tissues. Even though the adipocyte derived hormones, which are important factors in the cross-talk between tissues, appear to be the group of genes whose transcript levels are most affected in mutant mice, it is not possible at this time to determine a direct cause of the phenotype in mutant mice. For this a more detailed analysis of the mouse metabolome is necessary.

The putative role of Lsm16 in mRNA decapping does not provide insight into the cause of the reduction in size of the mutant mice. Since Lsm16 is believed to only enhance decapping, in mutant mice decapping would still be conducted. If a slower rate of decapping is taking place, this could lead to an apparent increase in transcript levels, as seen in the RT-PCR experiments. This is not supported though by the reduction in transcript levels seen in BAT and WAT. In addition, only 7 of the 21 genes analyzed showed changes in transcript levels in mutants. It is only possible to suggest that Lsm16 might be involved in enhancing decapping on a subset of mouse genes and when this function cannot be completely fulfilled it indirectly leads to a reduction in size of mutant mice compared to wild-types.

For the further analysis of the A008A01 mouse line it is necessary to verify the role of Lsm16 in mRNA decapping. Initial experiments should include not only decapping

assays to determine decapping activity, but also immunofluorescent staining to demonstrate the localization of Lsm16 to P-bodies. Antibodies against the decapping enzymes Dcp1 and Dcp2 could be used as markers for P-bodies.

4.4. Functional gene analysis by mutagenesis of the mouse genome

A long-term goal of molecular biology in the postgenomic era is the elucidation of the function of the estimated 30,000 genes in the mouse genome. To accomplish this, the gap between sequence information and knowledge of gene function will need to be closed. This will involve not only large-scale, genome-wide mutagenesis approaches, such as gene trap and ENU mutagenesis screens, but also gene specific approaches such as knock-outs, conditional mutagenesis, knock-ins, transgenic mice and RNAi.

Gene trapping represents a method for randomly mutating the mouse genome by insertional mutagenesis of embryonic stem cells (reviewed in: Gossler *et al*., 1989; Stanford *et al*., 2001; Floss and Wurst, 2002; Schnütgen *et al*., 2005). The main advantages of this method are suitability for high-throughput, genome-wide mutagenesis and ability to trap known, as well as novel genes. The randomness of the mutation means though, that there is no guarantee that any particular gene of interest will be trapped. Furthermore, multiple vector integrations per ES-cell can occur, though seldom. Such cells would become obvious through difficulties in determining the trapped sequence and thus not be considered for further analysis. Initial drawbacks to gene trapping were the inability to trap secretory genes, and those genes not expressed in ES-cells. These problems have been resolved by the development of new gene trap vectors, though no single vector type is able to trap all genes (Stanford *et al*., 2001; De-Zolt *et al*., 2006). The process of mutagenesis of the ES-cells and the determination and analysis of the trapped sequence lasts several months. At this step, gene trapping is of benefit for establishing mutagenized ES-cell libraries. The libraries allow for fast acquisition of not only a single particular gene, but also genes from families or processes of interest. As of November 2005, the International Gene Trap Consortium database contained 51244 ES-cell lines. Of these, 6347 were genes represented in the Ensembl database. This represents genome coverage of approximately 40.7%. From the analysis of 5142 sequence tags taken from gene-trap integrations, it was found that all functional classes of genes

can be disrupted. A random sampling of these insertions showed that 59% displayed an observable phenotype (Hansen *et al*., 2003). This is similar to the frequency obtained by conventional gene targeting. From an analysis of gene trap mice with observable phenotypes, it was found that 91% exactly phenocopied the gene targeted mutation (Stanford *et al*., 2001).

The most time consuming aspect of gene trapping is the analysis of the mutant mice. One of the first steps in the characterization is the determination of the spatio-temporal expression pattern of the trapped gene. Here, the *lacZ* reporter gene from the gene trap vector is of benefit. Because it is transcribed from the promoter of the trapped gene, the β-galactosidase activity of the fusion protein reflects, in general, the endogenous gene expression, though this can be affected by the loss of localization signals. Here, it is advantageous to use *in situ* hybridization to more accurately determine the expression of the endogenous gene (Skarnes *et al*., 1992). Further characterization of the mouse is based on detecting and analyzing the phenotype.

Since integration of the gene trap vector is random, it is possible that the vector integrates in the locus 3' to exons which encode functionally important domains. In this case, the mutation might not cause a null-mutation. This may lead to no obvious phenotype or to a hypomorph. However, even hypomorphs can be beneficial for analysis because partial gene inactivity is the bases for numerous human diseases. To minimize the time needed to analyze the gene trap mouse, a gene or genes can be selected that fall within the researcher's area of expertise. To further aid in the characterization, "mouse clinics" have been established which are designed for the analysis of any type of generated mouse mutant. These facilities offer large-scale standardized and comprehensive phenotypic analyses covering areas such as behavior, development, immunology, metabolism, organ function and pathology. With these screens, it is possible to detect less obvious phenotypes that would not be expected based on the knowledge of the trapped gene.

ENU mutagenesis represents a second common method for randomly mutating the mouse genome (de Angelis *et al*., 1998; Beier, 2000). Here, male mice are mutagenized. Their spermatocytes, which experience a high rate of mutation, carry

the mutation (Russell *et al*., 1979). The main advantage of this method is its suitability for large-scale, genome-wide mutagenesis. Here, the process of generating the mutant mice lasts several months. Due to the random nature of mutagenesis, several unavoidable disadvantages arise. These include multiple mutations per mouse, which can be eliminated by further breeding of the mouse, and incomplete inactivation of the gene, which can not be alleviated, though these mice would most likely not be detected in the phenotype screens (reviewed in Kennedy and O'Bryan, 2006). The most time consuming aspects of this methods are the time needed to identify the mutation, which can take up to one year, and the time to further analyze the mouse, which can take up to 2 years (Hafner and Müller, 2004). This method also requires extensive resources for establishing the large number of mice to be examined. For ENU, it is beneficiary and common to first set up screens to look for genes that share a phenotype or genes that affect the tissue, process or characteristic of interest (Justice, 2000). To further aid in the characterization, "mouse clinics" have been established.

Since gene trapping and ENU mutagenesis do not allow for the targeting of a desired gene, gene specific approaches are also a necessity for functional genome analysis in the mouse. The most common approach in this respect is the targeted gene replacement or knock-out. (Capecchi, 1994) The first step is to create a targeting vector, which after transfection into ES-cells, mutates the target locus by exchange of a stretch of the endogenous sequence with a stretch of altered sequence via homologous recombination. Although the sequence of the gene to be mutated is known, approximately twice the amount of time is needed to create the appropriate targeting vector and transfect ES-cells as compared to the time needed for the mutagenesis by gene trapping and ENU. In addition, the time required for identifying the ES-cells carrying the desired homologous recombination event is considerably longer than for the identification of successful insertion of the gene trap vector in ES-cells. The amount of time necessary to analyze the mouse is similar to that of gene trap mice. The knock-out, gene trap and ENU approaches share the common goal of mutating a gene to the extent that the protein produced is non- functional. Depending on the importance of the gene, this can lead to embryonic lethality in the mouse. This can be overcome by a variant of the knock-out method known as the conditional knock-out or conditional gene targeting approach.

Conditional gene targeting is similar to the traditional gene targeting or knock-out approach in that the desired sequence of the vector integrates by homologous recombination into the target locus in ES-cells. The ES-cells carrying the appropriate recombination event are then used to create the mutant mice. The main differences, which are also the advantages, of conditional knock-out over traditional knock-out are the type of vector used and the process of mouse analysis (Rajewsky *et al*., 1996; Sauer, 1998). The conditional targeting vector contains the portion of the endogenous sequence to be later removed, flanked by loxP sites or frt sites. These allow for the removal of the endogenous sequence by the Cre or FLP recombinase respectively, at a particular time point or in a certain cell type of interest. This is accomplished by mating the mice to transgenic mice expressing the recombinase at the time point or in the cell type desired, thus permitting a more sophisticated manipulation of the mouse genome. As disadvantages, the generation of the vector and the analysis of the mice require more time as compared to the traditional approach, and the number of mice needed is higher.

Another approach for disrupting gene function that is beginning to be utilized in the mouse is RNA interference (RNAi; Coumoul and Deng, 2006). This fast and simple method is based on the degradation of a specific target mRNA due to the presence of small perfectly homologous double-stranded RNAs, referred to as small interfering RNAs (siRNAs). This leads to a decrease or “knock-down” in expression of the corresponding protein. Because the expression of the target gene is not completely ablated, this approach is more appropriate as a complement to one of the methods mentioned above.

Gene function in the mouse can be analyzed not only by disrupting the gene, as seen above, but also by exogenous expression of the protein in some or all tissues. Two methods that use this approach are the transgenic mouse and the knock-in (reviewed in Babinet, 2000). In the case of transgenic mice, a new piece of DNA is introduced into the mouse genome. This piece of DNA includes the structural gene of interest, a strong mouse gene promoter and enhancer to be expressed and vector DNA to enable the transgene to be inserted into the mouse genome. Successful integration of this DNA results in the expression of the transgene in addition to the wild type, endogenous protein levels in the mouse. Depending on the goal of the experiment,

the transgenic mouse will exhibit over-expression of a non-mutated protein, expression of a dominant-negative form of a protein, or expression of a fluorescent-tagged protein. This method is relatively quick and simple, but includes the risk that the DNA may insert itself into a critical locus, causing an unexpected, detrimental genetic mutation. Alternatively, the transgene may insert into a locus that is subject to gene silencing. If the protein being expressed from the transgene causes toxicity, excessive over-expression from multiple insertions can be lethal to some tissues or even to the entire mouse. For these reasons, several independent mouse lines containing the same transgene must be created and studied to ensure that any resulting phenotype is not due to toxic gene-dosing or to the mutations created at the site of transgene insertion.

For analyzing the exogenous expression of a protein, the problems of a standard transgenic mouse, can be avoided by the knock-in approach. A knock-in mouse is generated by targeted insertion of the transgene at a selected locus. The insert is flanked by DNA from a non-critical locus, and homologous recombination allows the transgene to be targeted to that specific, non-critical integration site. In this way, the genetic environment surrounding the over-expression cassette can be controlled and it is likely that the DNA does not incorporate itself into multiple locations. Site-specific knock-ins result in a more consistent level of expression of the transgene from generation to generation because it is known that the over-expression cassette is present as a single copy. Also, because a targeted transgene is not interfering with a critical locus, it is more certain that any resulting phenotype is due to the exogenous expression of the protein. Although the generation of a knock-in mouse does avoid many of the problems of a traditional transgenic mouse, this procedure requires more time to assemble the vector and to identify ES-cells that have undergone homologous recombination.

Each method for analyzing gene function has its advantages and disadvantages. No single approach is ideal for all circumstances. It is therefore not a question of which method is the best in general, but instead, which method is the best for answering the question posed by the researcher. Since the final goal involves elucidating the function of all the genes in the mouse genome, especially those of biomedical relevance in humans, it will be necessary to utilize a wide range of mutagenesis technologies.

5. References

Agata, Y., Matsuda, E. and Shimizu, A. (1999). Two novel Krüppel-associated box-containing zinc-finger proteins, KRAZ1 and KRAZ2, repress transcription through functional interaction with the corepressor KAP-1 (TIF1ß/KRIP-1). *J Biol Chem* **274**, 16412-16422.

Albrecht, M. and Lengauer, T. (2004). Novel Sm-like proteins with long C-terminal tails and associated methyltransferases. *FEBS Letters* **569**, 18-26.

Anantharaman, V. and Aravind, L. (2004). Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability. *BMC Genomics* **5**, 45.

Babinet, C. (2000). Transgenic mice: an irreplaceable tool for the study of mammalian development and biology. *J Am Soc Nephrol* **11**, 88-94.

Beermann, F., Schmid, E. and Schutz, G. (1992). Expression of the mouse tyrosinase gene during embryonic development: recapitulation of the temporal regulation in transgenic mice. *PNAS* **89**, 2809-2813.

Beier, D. R. (2000). Sequenced-based analysis of mutagenized mice. *Mamm Genome* **11**, 594-597.

Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. and Martial, J. A. (1991). The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *PNAS* **88**, 3608-3612.

Beltzner, C. C. and Pollard, T. D. (2004). Identification of functionally important residues of Arp2/3 complex by analysis of homology models from diverse species. *J Mol Biol* **336**, 551-565.

Boldogh, I. R., Yang, H.-C., Nowakowski, W. D., Karmon, S. L., Hays, L. G., Yates, J. R., III and Pon, L. A. (2001). Arp2/3 complex and actin dynamics are required for actin-based mitochondrial motility in yeast. *PNAS* **98**, 3162-3167.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.

Brilliant, M. H., Ching, A., Nakatsu, Y. and Eicher, E. M. (1994). The original pink-eyed dilution mutation (p) arose in asiatic mice: implications for the H4 minor histocompatibility antigen, Myod1 regulation and the origin of inbred strains. *Genetics* **138**, 203-211.

Brilliant, M. H. (2001). The mouse p (pink-eyed dilution) and human P genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. *Pigment Cell Res* **14**, 86-93.

Brown, M. S., Ho, Y. K. and Goldstein, J. L. (1980). The cholesteryl ester cycle in macrophage foam cells. Continual hydrolysis and re-esterification of cytoplasmic cholesteryl esters. *J Biol Chem* **255**, 9344-9352.

Capecchi, M. R. (1994). Targeted gene replacement. *Sci Am* **270**, 52-59.

Chen, K., Manga, P. and Orlow, S. J. (2002). Pink-eyed dilution protein controls the processing of tyrosinase. *Mol Biol Cell* **13**, 1953-1964.

Chirala, S. S., Chang, H., Matzuk, M., Abu-Elheiga, L., Mao, J., Mahon, K., Finegold, M. and Wakil, S. J. (2003). Fatty acid synthesis is essential in embryonic development: fatty acid synthase null mutants and most of the heterozygotes die in utero. *PNAS* **100**, 6358-6363.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**, 156-159.

Coleman, R. A., Lewin, T. M. and Muoio, D. M. (2000). Physiological and nutritional regulation of enzymes of triacylglycerol synthesis. *Annu Rev Nutr* **20**, 77-103.

Cook, K. S., Groves, D. L., Min, H. Y. and Spiegelman, B. M. (1985). A developmentally regulated mRNA from 3T3 adipocytes encodes a novel serine protease homologue. *PNAS* **82**, 6480-6484.

Cooksey, C. J., Garratt, P. J., Land, E. J., Pavel, S., Ramsden, C. A., Riley, P. A. and Smit, N. P. M. (1997). Evidence of the indirect formation of the catecholic intermediate substrate responsible for the autoactivation kinetics of tyrosinase. *J Biol Chem* **272**, 26226-26235.

Cougot, N., Babajko, S. and Seraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. *J Cell Biol* **165**, 31-40.

Coumoul, X. and Deng, C.-X. (2006). RNAi in mice: a promising approach to decipher gene functions in vivo. *Biochimie* **In Press, Corrected Proof**.

Dai, K. S. and Liew, C. C. (1999). Chromosomal, in silico and in vitro expression analysis of cardiovascular-based genes encoding zinc finger proteins. *J Mol Cell Cardiol* **31**, 1749-1769.

De-Zolt, S., Schnutgen, F., Seisenberger, C., Hansen, J., Hollatz, M., Floss, T., Ruiz, P., Wurst, W. and von Melchner, H. (2006). High-throughput trapping of secretory pathway genes in mouse embryonic stem cells. *Nucl Acids Res* **34**, e25.

de Angelis, M. H. and Balling, R. (1998). Large scale ENU screens in the mouse: genetics meets genomics. *Mutat Res* **400**, 25-32.

Durham-Pierre, D., Gardner, J. M., Nakatsu, Y., King, R. A., Francke, U., Ching, A., Aquaron, R., del Marmol, V. and Brilliant, M. H. (1994). African origin of an intragenic deletion of the human P gene in tyrosinase positive oculocutaneous albinism. *Nat Genet* **7**, 176-179.

Engqvist-Goldstein, A. E. Y. and Drubin, D. G. (2003). Actin assembly and endocytosis: from yeast to mammals. *Ann Rev Cell Dev Biol* **19**, 287-332.

Fenger-Gron, M., Fillman, C., Norrild, B. and Lykke-Andersen, J. (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol Cell* **20**, 905-915.

Flach, G., Johnson, M. H., Braude, P. R., Taylor, R. A. and Bolton, V. N. (1982). The transition from maternal to embryonic control in the 2-cell mouse embryo. *EMBO J* **1**, 681-686.

Fleming, T. P., Butler, E., Collins, J., Sheth, B. and Wild, A. E. (1998). Cell polarity and mouse early development. *Adv Mol Cell Biol* **26**, 67-94.

Floss, T. and Wurst, W. (2002). Functional genomics by gene-trapping in embryonic stem cells. *Methods Mol Biol* **185**, 347-379.

Freese, E. (2003). Charackterisierung des Arp3 - Gens in der Maus. *Diplomarbeit, TU Braunschweig*.

Friedman, J. M. (2003). A war on obesity, not the obese. *Science* **299**, 856-858.

Friedrich, G. and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5**, 1513-1523.

Fu, Y., Luo, N., Klein, R. L. and Garvey, W. T. (2005). Adiponectin promotes adipocyte differentiation, insulin sensitivity, and lipid accumulation. *J Lipid Res* **46**, 1369-1379.

Garreis, F. (2005). Charakterisierung der mutanten Mauslinien A008A01^{GT} und Ataxin-2^{GT}. *Diplomarbeit, TU Braunschweig*.

Gossler, A., Joyner, A. L., Rossant, J. and Skarnes, W. C. (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* **244**, 463-465.

Gouin, E., Welch, M. D. and Cossart, P. (2005). Actin-based motility of intracellular pathogens. *Curr Opin Microbiol* **8**, 35-45.

Hafner, M. and Müller, W. (2004). Generation of mouse mutants by sequence information driven and random mutagenesis. In *The Laboratory Mouse*, (ed. H. J. Hedrich), pp. 85-95. Amsterdam: Elsevier.

Halaban, R., Moellmann, G., Tamura, A., Kwon, B. S., Kuklinska, E., Pomerantz, S. H. and Lerner, A. B. (1988). Tyrosinases of murine melanocytes with mutations at the albino locus. *PNAS* **85**, 7241-7245.

Hammond, L. E., Gallagher, P. A., Wang, S., Hiller, S., Kluckman, K. D., Posey-Marcos, E. L., Maeda, N. and Coleman, R. A. (2002). Mitochondrial glycerol-3-phosphate acyltransferase-deficient mice have reduced weight and liver triacylglycerol content and altered glycerolipid fatty acid composition. *Mol Cell Biol* **22**, 8204-8214.

Hansen, J., Floss, T., Van Sloun, P., Fuchtbauer, E.-M., Vauti, F., Arnold, H.-H., Schnutgen, F., Wurst, W., von Melchner, H. and Ruiz, P. (2003). A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *PNAS* **100**, 9918-9922.

Harada, K., Shen, W.-J., Patel, S., Natu, V., Wang, J., Osuga, J.-I., Ishibashi, S. and Kraemer, F. B. (2003). Resistance to high-fat diet-induced obesity and altered expression of adipose-specific genes in HSL-deficient mice. *Am J Physiol Endocrinol Metab* **285**, E1182-E1195.

He, W. and Parker, R. (2000). Functions of Lsm proteins in mRNA degradation and splicing. *Curr Opin Cell Biol* **12**, 346-350.

Hermann, H., Fabrizio, P., Raker, V. A., Foulaki, K., Hornig, H., Brahms, H. and Lührmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* **14**, 2076-2088.

Hudson, A. M. and Cooley, L. (2002). A subset of dynamic actin rearrangements in Drosophila requires the Arp2/3 complex. *J Cell Biol* **156**, 677-687.

Jheon, A. H., Ganss, B., Cheifetz, S. and Sodek, J. (2001). Characterization of a novel KRAB/C2H2 zinc finger transcription factor involved in bone development. *J Biol Chem* **276**, 18282-18289.

Johnson, M. H., Maro, B. and Takeichi, M. (1986). The role of cell adhesion in the synchronisation and orientation of polarization in 8-cell mouse blastomeres. *J Embryol Exp Morph* **93**, 239-255.

Justice, M. J. (2000). Capitalizing on large-scale mouse mutagenesis screens. *Nat Rev Genet* **1**, 109-115.

Kennedy, C. L. and O'Bryan, M. K. (2006). N-ethyl-N-nitrosourea (ENU) mutagenesis and male fertility research. *Hum Reprod Update*, 1-9.

Khusial, P., Plaag, R. and Zieve, G. W. (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem Sci* **30**, 522-528.

Kim, J. B., Sarraf, P., Wright, M., Yao, K. M., Mueller, E., Solanes, G., Lowell, B. B. and Spiegelman, B. M. (1998). Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. *J Clin Invest* **101**, 1-9.

King, R. A., Lewis, R. A., Townsend, D., Zelickson, A., Olds, D. P. and Brumbaugh, J. (1985). Brown oculocutaneous albinism: clinical, ophthalmological, and biochemical characterization. *Ophthal* **92**, 1496-1505.

Klug, A. and Schwabe, J. W. (1995). Protein motifs 5, zinc finger. *FASEB J* **9**, 597-604.

Kovacs, E. M., Goodwin, M., Ali, R. G., Paterson, A. D. and Yap, A. S. (2002). Cadherin-directed actin assembly: E-cadherin physically associates with the Arp2/3 complex to direct actin assembly in nascent adhesive contacts. *Curr Biol* **12**, 379-382.

Krebs, C. J., Larkins, L. K., Khan, S. M. and Robins, D. M. (2005). Expansion and diversification of KRAB zinc-finger genes within a cluster including Regulator of sex-limitation 1 and 2. *Genomics* **85**, 752-761.

Kshirsagar, M. and Parker, R. (2004). Identification of Edc3p as an enhancer of mRNA decapping in saccharomyces cerevisiae. *Genetics* **166**, 729-739.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lamoreux, M. L., Wakamatsu, K. and Ito, S. (2001). Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. *Pigment Cell Res* **14**, 1-23.

Lamoreux, M. L., Zhou, B. K., Rosemblat, S. and Orlow, S. J. (1995). The pinkeyed-dilution protein and the eumelanin/pheomelanin switch: in support of a unifying hypothesis. *Pigment Cell Res* **5**, 263-270.

Larue, L., Ohsugi, M., Hirchenhain, J. and Kemler, R. (1994). E-Cadherin null nutant embryos fail to form a trophectoderm epithelium. *PNAS* **91**, 8263-8267.

Lazar, M. A. (2005). How obesity causes diabetes: not a tall tale. *Science* **307**, 373-379.

Le Fur, N., Kelsall, S. R., Silvers, W. K. and Mintz, B. (1997). Selective increase in specific alternative splice variants of tyrosinase in murine melanomas: A projected basis for immunotherapy. *PNAS* **94**, 5332-5337.

Le, J., El-Assal, S. E.-D., Basu, D., Saad, M. E. and Szymanski, D. B. (2003). Requirements for Arabidopsis ATARP2 and ATARP3 during epidermal development. *Curr Biol* **13**, 1341-1347.

Lehtonen, E. and Badley, R. A. (1980). Localization of cytoskeletal proteins in preimplantation mouse embryos. *J Embryol Exp Morph* **55**, 211-225.

Li, Q., Li, Z., Sun, C.-X. and Yu, A. C.-H. (2002). Identification of transcripts expressed under functional differentiation in primary culture of cerebral cortical neurons. *Neurochem Res* **27**, 147-154.

Machesky, L. M., Atkinson, S. J., Ampe, C., Vandekerckhove, J. and Pollard, T. D. (1994). Purification of a cortical complex containing two unconventional actins from Acanthamoeba by affinity chromatography on profilin-agarose. *J Cell Biol* **127**, 107-115.

Machesky, L. M., Mullins, R. D., Higgs, H. N., Kaiser, D. A., Blanchoin, L., May, R. C., Hall, M. E. and Pollard, T. D. (1999). Scar, a WASp-related protein, activates nucleation of actin filaments by the Arp2/3 complex. *PNAS* **96**, 3739-3744.

Machesky, L. M., Reeves, E., Wientjes, F., Mattheyse, F. J., Grogan, A., Totty, N. F., Burlingame, A. L., Hsuan, J. J. and Segal, A. W. (1997). Mammalian actin-related protein 2/3 complex localizes to regions of lamellipodial protrusion and is composed of evolutionarily conserved proteins. *Biochem J* **328**, 105-112.

Margolin, J. F., Friedman, J. R., Meyer, W. K.-H., Vissing, H., Thiesen, H.-J. and Rauscher III, F. J. (1994). Krüppel-associated boxes are potent transcriptional repression domains. *PNAS* **91**, 4509-4513.

Mark, C., Looman, C., Abrink, M. and Hellman, L. (2001). Molecular cloning and preliminary functional analysis of two novel human KRAB zinc finger proteins, HKr18 and HKr19. *DNA Cell Biol* **20**, 275-286.

Merrifield, C. J., Qualmann, B., Kessels, M. M. and Almers, W. (2004). Neural Wiskott Aldrich Syndrome Protein (N-WASP) and the Arp2/3 complex are recruited to sites of clathrin-mediated endocytosis in cultured fibroblasts. *Eur J Cell Biol* **83**, 13-18.

Meyer, N. (2003). Charakterisierung eines neuen Gens in der Maus mittels Gene-Trap-Mutagenese. *Diplomarbeit, TU Braunschweig*.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.-P. and Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1ß that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res* **24**, 4859-4867.

Moreau, V., Madania, A., Martin, R. P. and Winson, B. (1996). The Saccharomyces cerevisiae actin-related protein Arp2 is involved in the actin cytoskeleton. *J Cell Biol* **134**, 117-132.

Mullins, R. D., Heuser, J. A. and Pollard, T. D. (1998). The interaction of Arp2/3 complex with actin: Nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *PNAS* **95**, 6181-6186.

Nagafuchi, A. and Takeichi, M. (1989). Transmembrane control of cadherin-mediated cell adhesions: a 94 kDa protein functionally associated with a specific region of the cytoplasmic domain of E-cadherin. *Cell Regul* **1**, 37-44.

Oetting, W. S. (2000). The tyrosinase gene and oculocutaneous albinism type 1 (OCA1): a model for understanding the molecular biology of melanin formation. *Pigment Cell Res* **13**, 320-325.

Ogilvie, R. W. and Feeback, D. L. (1990). A metachromatic dye-ATPase method for the simultaneous identification of skeletal muscle fiber types I, IIA, IIB and IIC. *Stain Technology* **65**, 231-241.

Ozawa, M., Baribault, H. and Kemler, R. (1989). The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J* **6**, 1711-1717.

Paynton, B. V., Rempel, R. and Bachvarova, R. (1988). Changes in state of adenylation and time course of degradation of maternal mRNAs during oocyte maturation and early embryonic development in the mouse. *Dev Biol* **129**, 304-314.

Piko, L. and Clegg, K. B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Dev Biol* **89**, 362-378.

Pollard, T. D. and Beltzner, C. C. (2002). Structure and function of the Arp2/3 complex. *Curr Opin Struct Biol* **12**, 768-774.

Porter, S. and Mintz, B. (1991). Multiple alternatively spliced transcripts of the mouse tyrosinase-encoding gene. *Gene* **97**, 277-282.

Prochnow, B. (2002). Mutagenesis of mouse embryonic stem cells by the gene trap approach: analysis of a new mouse line with a gene trap vector integration in a novel gene. *Diploma Thesis, TU Braunschweig*.

Puri, N., Gardner, J. M. and Brilliant, M. H. (2000). Aberrant pH of melanosomes in pink-eyed dilution (p) mutant melanocytes. *J Invest Dermatol* **115**, 607-613.

Qi, Y., Takahashi, N., Hileman, S. M., Patel, H. R., Berg, A. H., Pajvani, U. B., Scherer, P. E. and Ahima, R. S. (2004). Adiponectin acts in the brain to decrease body weight. *Nat Med* **10**, 524-529.

Ramsay, M., Colman, M. A., Stevens, G., Zwane, E., Kromberg, J., Farrall, M. and Jenkins, T. (1992). The tyrosinase-positive oculocutaneous albinism locus maps to chromosome 15q11.2-q12. *Am J Hum Genet* **51**, 879-884.

Rajewsky, K., Gu, H., Kuhn, R., Betz, U. A. K., Muller, W., Roes, J. and Schwenk, F. (1996). Conditional gene targeting. *J Clin Invest* **98**, 600-603.

Rangwala, S. M., Rich, A. S., Rhoades, B., Shapiro, J. S., Obici, S., Rossetti, L. and Lazar, M. A. (2004). Abnormal glucose homeostasis due to chronic hyperresistinemia. *Diabetes* **53**, 1937-1941.

Riethmacher, D., Brinkmann, V. and Birchmeier, C. (1995). A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. *PNAS* **92**, 855-859.

Rinchik, E. M., Bultman, S. J., Horsthemke, B., Lee, S.-T., Strunk, K. M., Spritz, R. A., Avidano, K. M., Jong, M. T. C. and Nicholls, R. D. (1993). A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. *Nature* **361**, 72-76.

Robinson, R. C., Turbedsky, K., Kaiser, D. A., Marchand, J.-B., Higgs, H. N., Choe, S. and Pollard, T. D. (2001). Crystal structure of Arp2/3 complex. *Science* **294**, 1679-1684.

Rosen, B. S., Cook, K. S., Yaglom, J., Groves, D. L., Volanakis, J. E., Damm, D., White, T. and Spiegelman, B. M. (1989). Adipsin and complement factor D activity: an immune-related defect in obesity. *Science* **244**, 1483-1487.

Rosemblat, S., Durham-Pierre, D., Gardner, J. M., Nakatsu, Y., Brilliant, M. H. and Orlow, S. J. (1994). Identification of a melanosomal membrane protein encoded by the pink-eyed dilution (type II oculocutaneous albinism) gene. *PNAS* **91**, 12071-12075.

Sawa, M., Suetsugu, S., Sugimoto, A., Miki, H., Yamamoto, M. and Takenawa, T. (2003). Essential role of the C. elegans Arp2/3 complex in cell migration during ventral enclosure. *J Cell Sci* **116**, 1505-1518.

Schnabl, B., Hu, K., Muhlbauer, M., Hellerbrand, C., Stefanovic, B., Brenner, D. A. and Scholmerich, J. (2005). Zinc finger protein 267 is up-regulated during the activation process of human hepatic stellate cells and functions as a negative transcriptional regulator of MMP-10. *Biochem Biophys Res Comm* **335**, 87.

Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. and Rauscher, F. J., III. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16**, 919-932.

Schultz, R. M. (1993). Regulation of zygotic gene activation in the mouse. *Bioessays* **8**, 531-538.

Seraphin, B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J* **14**, 2089-2098.

Shulman, D. I. (2000). Metabolic effects of growth hormone in the child and adolescent. *Curr Opin Pediatr* **4**, 432-436.

Silvers, W. K. (1979). The coat colors of mice. A model for mammalian gene action and interaction. Springer, New York.

Simpson, E. M., Linder, C. C., Sargent, E. E., Davisson, M. T., Mobraaten, L. E. and Sharp, J. J. (1997). Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat Genet* **16**, 19-27.

Skarnes, W. C., Auerbach, A. and Joyner, A. L. (1992). A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev* **6**, 903-918.

Stanford, W. L., Cohn, J. B. and Cordes, S. P. (2001). Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2**, 756-768.

Steppan, C. M., Bailey, S. T., Bhat, S., Brown, E. J., Banerjee, R. R., Wright, C. M., Patel, H. R., Ahima, R. S. and Lazar, M. A. (2001). The hormone resistin links obesity to diabetes. *Nature* **409**, 307-312.

Taunton, J., Rowning, B. A., Coughlin, M. L., Wu, M., Moon, R. T., Mitchison, T. J. and Larabell, C. A. (2000). Actin-dependent propulsion of endosomes and lysosomes by recruitment of N-WASP. *J Cell Biol* **148**, 519-530.

Terao, M., Tabe, L., Garattini, E., Sartori, D., Studer, M. and Mintz, B. (1989). Isolation and characterization of variant cDNAs encoding mouse tyrosinase. *Biochem Biophys Res Comm* **2**, 848-853.

Threadgill, D. W., Yee, D., Matin, A., Nadeau, J. H. and Magnuson, T. (1997). Genealogy of the 129 inbred strains: 129/SvJ is a contaminated inbred strain. *Mamm Genome* **8**, 390-393.

Torres, M., Stoykova, A., Huber, O., Chowdhury, K., Bonaldo, P., Mansouri, A., Butz, S., Kemler, R. and Gruss, P. (1997). An alpha -E-catenin gene trap mutation defines its function in preimplantation development. *PNAS* **94**, 901-906.

Totsuka, Y., Nagao, Y., Horii, T., Yonekawa, H., Imai, H., Hatta, H., Izaike, Y., Tokunaga, T. and Atomi, Y. (2003). Physical performance and soleus muscle fiber composition in wild-derived and laboratory inbred mouse strains. *J Appl Physiol* **95**, 720-727.

Truett, G. E., Heeger, P., Mynatt, R. L., Truett, A. A., Walker, J. A. and Warman, M. L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **1**, 52-54.

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. *Genome Biol* **4**, 231.

Verma, S., Shewan, A. M., Scott, J. A., Helwani, F. M., Elzen, N. R. d., Miki, H., Takenawa, T. and Yap, A. S. (2004). Arp2/3 activity is necessary for efficient formation of E-cadherin adhesive contacts. *J Biol Chem* **279**, 34062-34070.

Wakil, S. J. (1989). Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**, 4523-4530.

Wang, Y.-X., Zhang, C.-L., Yu, R. T., Cho, H. K., Nelson, M. C., Bayuga-Ocampo, C. R., Ham, J., Kang, H. and Evans, R. M. (2004). Regulation of muscle fiber type and running endurance by PPARg. *PLoS Biology* **2**, e294.

Welch, M. D., DePace, A. H., Verma, S., Iwamatsu, A. and Mitchison, T. J. (1997). The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly. *J Cell Biol* **138**, 375-384.

Welch, M. D. and Mullins, R. D. (2002). Cellular control of actin nucleation. *Annu Rev Cell Dev Biol* **18**, 247-288.

Witkop, C. J. J., Niswander, J. D., Bergsma, D. R., Workman, P. L. and White, J. G. (1972). Tyrosinase positive oculocutaneous albinism among the Zuni and the Brandywine triracial isolate: biochemical and clinical characteristics and fertility. *Am J Phys Anthropol* **36**, 397-405.

Witzgall, R., O'Leary, E., Leaf, A., Önaldi, D. and Bonventre, J. V. (1994). The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *PNAS* **91**, 4514-4518.

Wong, G. W., Wang, J., Hug, C., Tsao, T.-S. and Lodish, H. F. (2004). A family of Acrp30/adiponectin structural and functional paralogs. *PNAS* **101**, 10302-10307.

Yarar, D., Waterman-Storer, C. M. and Schmid, S. L. (2005). A dynamic actin cytoskeleton functions at multiple stages of clathrin-mediated endocytosis. *Mol Biol Cell* **16**, 964-975.

Yu, Y.-H. and Ginsberg, H. N. (2005). Adipocyte signaling and lipid homeostasis: sequelae of insulin-resistant adipose tissue. *Circ Res* **96**, 1042-1052.

6. Abbreviations

aa: Amino acid

AEBSF: 4-(2-Aminoethyl) benzenesulphonyl fluoride

Arp: Actin-related protein

BAT: Brown adiopose tissue

βgal: β-galactosidase

βgeo: β-galactosidase/neomycin-phosphotransferase fusion gene

BLAST: Basic local alignment search tool

Bluo-gal: 5-bromo-3-indolyl-beta-D-galactopyranoside

bp: Base pair

BSA: Bovine serum albumin

C-terminus: Carboxy-terminus

CDART: Conserved domain architecture retrieval tool

cDNA: Complementary DNA

DAB: 3,3´-Diaminobenzidine

DEPC: Diethyl pyrocarbonate

DMF: Dimethylformamide

DOC: Sodium desoxycholate

DNA: Deoxyribonucleic acid

dNTP: Deoxyribonucleoside triphosphate

ECL: Enhanced chemiluminescence

EDTA: Ethylene diamine tetraacetate

EGTA: Ethylene glycol bis-(2-aminoethyl ether)tetraacetate

ENU: N-ethyl-N-nitrosourea

ES-cell: Embryonic stem cell

EST: Expressed sequence tag

GGTC: German gene trap consortium

H+L: Antibody heavy and light chains

HRP: Horseradish peroxidase

Ig: Immunoglobulin

IMAGE: Integrated molecular analysis of genomes and their expression

kb: Kilobase pair

kDa: Kilodalton

KRAB: Krüppel-associated box

LacZ: β-galactosidase

Lsm: like-Sm

MOPS: 3-[N-Morpholino] propanesulfonic acid

mRNA: Messenger RNA

N-terminus: Amino-terminus

NCBI: National center for biotechnology information

neo: Neomycin-phosphotransferase

nt: Nucleotide

OD: Optical density

ORF: Open reading frame

PBS: Phosphate-buffered saline

PFA: Paraformaldehyde

PCR: Polymerase Chain Reaction

Pfam: Protein families database

RACE: Rapid amplification of cDNA ends

RNA: Ribonucleic acid

RT-PCR: Reverse transcription PCR

SA: Splice-acceptor

SVpA: Simian virus polyadenylation signal

TBST: Tris-buffered saline tween-20

TBSX: Tris-buffered saline triton X-100

Tris: Tris-(hydroxymethyl) aminomethane

UV: Ultraviolet

WAT: White adipose tissue

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactoside

ZF: Zinc-finger

Acknowledgements

I would like to thank Prof. H.-H. Arnold for providing me with three interesting gene trap mouse lines to analyze as the basis for my thesis, as well as for the opportunity to work in the Department of Cell and Molecular Biology on a Georg-Christoph-Lichtenberg Fellowship.

My gratitude also goes to my supervisor Dr. Franz Vauti for the support and guidance during my work in the lab.

I am thankful to Prof. Dr. B. M. Jockusch not only for her willingness to grade my thesis and take-part on my examination committee, but also for taking the time to talk non-science with me.

To Prof. Dr. M. Korte I am grateful for his acceptance to act as chairman of my examination committee.

I am very appreciative of Dr. Björg Pauling for always finding the answers to my IGC questions and of Charlotte Klaue for not letting bureaucratic paper-work get in the way of her helping me.

Many thanks go to my lab mates present and past: Friederike Kruse, Sigi Düerkop, Monika Giesen, Annika Krause and Stefanie Willenzon for keeping the lab in working order, Nicole Meyer, Tobias Goller, Mandy Reichenbach, Elke Freese, Fabian Garreis and Christine Deller for always creating a friendly and definitely interesting lab atmosphere.

I also wish to thank the rest of my co-workers in the Department of Cell and Molecular Biology, especially Dr. Heike Kollmus for their answers to my many questions over the years.

A special thanks goes to the animal caretakers Irina Tunger, Claudia Schultze and Birgit Römmel for the day-in and day-out work of looking after my mice, especially when I couldn't find the time to make sure everything was going fine.

For their support, patience and understanding over the last four years, I am most thankful to my son Alexander, my wife Conni and her parents. If they had been qualified, they would have been standing next to me in the lab helping me with my experiments. I am also greatly thankful to my family in the USA, even though they were an ocean away, I knew they wished they could have been right here with me.